Neurodevelopmental outcome of children born following assisted reproductive technology

Middelburg, Karin

NEURODEVELOPMENTAL OUTCOME OF CHILDREN BORN FOLLOWING ASSISTED REPRODUCTIVE TECHNOLOGY

0 - 2 YEARS

The studies in this thesis have been carried out with financial support from the University Medical Center Groningen (grant 754510), the Dutch Organisation for Health Research and Development, (ZonMw grant 945-03-013 follow-up), the Research School of Behavioural and Cognitive Neurosciences and the Cornelia Stichting.

Publication of this thesis was financially supported by the University of Groningen, University Medical Center Groningen, Origio Benelux B.V., Goodlife Fertility B.V., Ferring B.V., Memidis Pharma B.V., MDT Artworks B.V. and Merck Sharp & Dome B.V..

Their support is gratefully acknowledged.

ISBN: 978-90-367-5096-7

NETZODRUK
DIGITAAL DRUKWERK

Cover and lay-out by Michiel Mellens, NetzoDruk Groningen.
Printed by NetzoDruk, Groningen.

rijksuniversiteit
groningen

NEURODEVELOPMENTAL OUTCOME OF CHILDREN BORN FOLLOWING ASSISTED REPRODUCTIVE TECHNOLOGY

0 - 2 YEARS

Proefschrift

ter verkrijging van het doctoraat in de
Medische Wetenschappen
aan de Rijksuniversiteit Groningen
op gezag van de
Rector Magnificus, dr. E. Sterken,
in het openbaar te verdedigen op
woensdag 28 september 2011
om 16.15 uur

door

Karin Janette Middelburg
geboren op 24 juli 1979
te Assen

Promotores

Prof. dr. M. Hadders-Algra
Prof. dr. M.J. Heineman
Prof. dr. A.F. Bos

Beoordelingscommissie

Prof. dr. F.M. Helmerhorst
Prof. dr. J.H. Kok
Prof. dr. C.M. Bilardo

CONTENTS

endless forms
most beautiful and most wonderful
have been,
and are being evolved
from so simple a beginning

adapted from Charles Darwin,
Origin of Species, 1859

Voor mijn ouders

1 General Introduction

BACKGROUND

The number of children born following assisted reproductive technology (ART) has shown a steady rise during the last decades. In 1978, Steptoe and Edwards reported the first pregnancy and live birth following *in vitro* fertilisation (IVF) (Steptoe and Edwards, 1978). Since then, numbers increased substantially. In 2010, Edwards was even awarded a Nobel Price, as his achievements in the development of ART are considered to represent a milestone in modern medicine. Meanwhile, up to 4% of children are born following ART in several European countries and it is likely that in the future this percentage will rise even further (Nyboe Andersen *et al.*, 2009).

As a consequence of the growing number of ART-children, health and development of these children has become of general importance. Over the years, it is well established that ART-children have an increased risk of being born preterm or with a low birth weight compared to their naturally conceived peers (Australian in vitro fertilisation collaborative group, 1985; Helmerhorst *et al.*, 2004; Jackson *et al.*, 2004). This difference is partly the result of the relatively high incidence of twins among ART-children. Yet, poorer outcomes are also found in singleton pregnancies (Australian in vitro fertilisation collaborative group, 1985; Helmerhorst *et al.*, 2004; Jackson *et al.*, 2004). The fact that preterm birth and low birth weight are strongly related to impaired neurodevelopment (Bhutta *et al.*, 2002) and diseases in adult life (Barker, 1995) necessitates careful follow-up of ART-children.

In this thesis the neurodevelopmental outcome of ART-children is evaluated up to the age of two years. An attempt is made to unravel the biological mechanisms that may underlie potentially poorer neurodevelopmental outcome. To clarify which mechanisms may be involved, I will shortly address the techniques used in conventional IVF, IVF in the modified natural cycle and IVF with preimplantation genetic screening.

Assisted Reproductive Technology

Since its introduction, ART has undergone several technological developments. These refinements were, in general, aimed at reducing side effects and improving success rate.

Conventional IVF consists of the following phases. First, the woman's monthly hormone cycle is suppressed with drugs (Gonadotrophine Releasing Hormone (GnRH)-agonists or antagonists) in order to avoid a Luteinising Hormone (LH)-surge, and untimely ovulation. In addition, the ovaries are stimulated to develop multiple follicles by administration of Follicle Stimulating Hormone (FSH). Growth of the follicles is carefully monitored with regular ultrasounds and measurement of estradiol levels. When several follicles have reached diameters of

approximately 18-20 mms and estradiol levels are sufficiently high, maturation of the follicles is boosted by supplementation of human Chorionic Gonadotrophin (hCG). Oocyte retrieval, by transvaginal ultrasound-guided follicle aspiration, is planned approximately 34 hours later. In the laboratory, the retrieved oocytes are then inseminated either naturally - by a selected sperm sample - or with intracytoplasmic sperm injection (ICSI). In the latter procedure, one spermatozoon is selected on its appearance and injected into the oocytes' cytoplasm (Palermo *et al.*, 1992). After *in vitro* fertilisation, the zygotes are cultured for 2 to 3 days. During this period, several cell cleavages take place, resulting in the origination of an early embryo. Hereafter, one or two of the morphologically best-looking embryos are transferred to the uterus. In case a woman achieves pregnancy, the endometrium may be supported by supplementation of progesterone or hCG. Surplus good-looking embryos may be cryopreserved and transferred in subsequent cycles if pregnancy does not occur (Zeilmaker *et al.*, 1984).

IVF in the Modified Natural Cycle

An alternative to conventional IVF is IVF in the modified natural cycle (MNC-IVF). The aim of this procedure is to use the one follicle that naturally develops to dominance (Nargund *et al.*, 2007). Like in conventional IVF, follicle size and estradiol levels are monitored closely. In contrast to conventional IVF, only when the lead follicle has reached a diameter of approximately 14 mms, GnRH- antagonists and FSH are started, mostly resulting in maturation of one single follicle (Rongières-Bertrand *et al.*, 1999;Pelinck *et al.*, 2005). With this approach, medication use is strongly reduced and the natural selection method for the dominant follicle is preserved. A resulting limitation, however, is a reduced pregnancy rate per cycle. Advantages of MNC-IVF above conventional IVF are a close to zero risk of multiple gestation and a negligible risk of ovarian hyperstimulation syndrome (Pelinck *et al.*, 2007).

Preimplantation Genetic Screening

The intention of IVF with preimplantation genetic screening (PGS) is to enhance the efficiency of assisted reproduction. An increase in numerical chromosomal abnormalities (aneuploidies) in ageing women may be one of the causes of a decrease in the chance of pregnancy with age (Wilton, 2002). Therefore, the concept of PGS is to identify and discard embryos with an abnormal chromosomal constitution. In the procedure, a hole is made in the zona pellucida of an embryo with laser or by chemical means. Subsequently, one or two blastomeres are aspirated so that copy numbers of several chromosomes can be determined with fluorescence in situ hybridisation (FISH). Theoretically, PGS should lead to higher ongoing pregnancy rates, however, in 2007 a randomised controlled trial on the

efficiency of PGS showed reduced instead of improved pregnancy rates following PGS (Mastenbroek *et al.*, 2007). This finding was confirmed in a recent meta-analysis (Mastenbroek *et al.*, 2008). For this reason, the technique, as described above, is no longer practiced on routine basis. Currently, feasibility of alternative forms of PGS (such as polar body biopsy and more comprehensive chromosome testing) is studied (Harper *et al.*, 2010;Geraedts *et al.*, 2010).

Mechanisms that may underlie poorer outcome following ART

It is conceivable that one or more components of the ART-procedure induce alterations in embryo development and in this way influence health and development of children born following ART. So far, the majority of studies focussed on the relation between ART and birth weight or other perinatal outcomes. From animal studies, it is well known that culture conditions affect birth weight, likely because of disturbed genomic imprinting in early embryo development (Young *et al.*, 2001;Khosla *et al.*, 2001;Ceelen and Vermeiden, 2001). In addition, hormonal hyperstimulation is proven to affect birth weight in animal models (Ertzeid and Storeng, 2001;Ceelen and Vermeiden, 2001;van der Auwera and D'Hooghe, 2001). Explanations that have been postulated for less optimal perinatal outcome are impaired endometrial receptivity because of altered hormone levels in ART or loss of natural selection of the dominant follicle, resulting in reduced oocyte quality (Ertzeid and Storeng, 2001;van der Auwera and D'Hooghe, 2001;Pelinck *et al.*, 2010). In studies concerning humans, type of culture medium (Dumoulin *et al.*, 2010) as well as ovarian hyperstimulation (Olivennes *et al.*, 1993;Ombelet *et al.*, 2006;Klemetti *et al.*, 2010) seem to affect birth weight. Remarkably, children born following frozen-thawed embryo transfer show higher birth weights and gestational ages than children born following fresh embryo transfer (Källén *et al.*, 2005;Pinborg *et al.*, 2009). Possibly, this is explained by transfer of the former group in naturally unstimulated cycles instead of hyperstimulated cycles with supra-physiological estradiol levels, which may effect endometrial receptivity. Alternatively, higher birth weights after freezing and thawing may be caused by selection of superior embryos by the cryopreservation procedure (Pinborg *et al.*, 2009; Källén *et al.*, 2005). However, since a recent study reported no linear relation between parameters of ovarian stimulation, such as doses of gonadotrophins used or duration of stimulation and birth weight (Griesinger *et al.*, 2008), it remains unclear to what extent ovarian hyperstimulation affects human birth weight.

In humans, subfertility itself is another factor suggested to be an important contributor to outcome of ART children. Subfertile couples are known to have an increased risk of obstetric complications and adverse perinatal outcome, including

increased risks of preeclampsia, antepartum haemorrhage, caesarean section, preterm birth, low birth weight and perinatal death (Thomson *et al.*, 2005;Pandian *et al.*, 2001;Draper *et al.*, 1999). Since some studies (Kapiteijn *et al.*, 2006;De Geyter *et al.*, 2006), but not others (Romundstad *et al.*, 2008) reported worse perinatal outcome in ART-children when compared to children of subfertile couples, it is still difficult to determine to what extent subfertility explains worse perinatal outcome following ART. Truly unravelling the effects of ART and the underlying indication for treatment can only be done by means of a trial with random allocation of assisted and natural conception, which is unethical and therefore impossible (Buck Louis *et al.*, 2005;Knoester, 2007).

Finally, vanishing twins may affect outcome following ART-pregnancies. Singleton pregnancies after ART are often the result of spontaneously reduced twin pregnancies and these so-called vanishing twins are related to preterm birth and low birth weight (Dickey *et al.*, 2002;Pinborg *et al.*, 2005;Pinborg *et al.*, 2007).

The above described factors that may influence perinatal outcome following ART may also influence development and health of ART-children. It would be of great interest to elucidate the mechanisms that underlie these associations. This could lead to the identification of at-risk subfertile couples and the provision of customised fertility, obstetrical or child-welfare care (Thomson *et al.*, 2005).

AIMS

The present thesis describes two projects. The aim of the first project is to disentangle the potential effects of controlled ovarian hyperstimulation and the *in vitro* procedure itself on neurodevelopmental outcome in infancy. In the second project, the aim is to study whether PGS in addition to 'conventional' IVF or ICSI affects neurodevelopmental outcome in infancy.

METHODS

Two separate projects are described in the present thesis. First, studies resulting from the data of the Groningen ART-cohort and secondly, studies on follow-up of children born after IVF with PGS. The two projects ran in parallel.

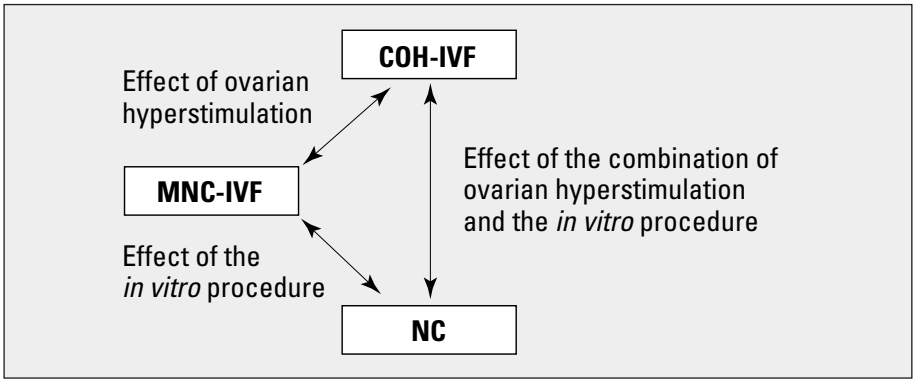
The Groningen ART-cohort study

The Groningen ART-cohort consists of three groups of prospectively recruited singletons. Couples were recruited at the Department of Reproductive Medicine of

the University Medical Center Groningen, a tertiary care centre in the Netherlands. All couples who were pregnant after IVF or ICSI, with a term date between March 2005 and December 2006 were asked to participate in a longitudinal study on neurodevelopmental outcome of IVF/ICSI children, during the third trimester of their pregnancy. This resulted in two groups; children born after 'conventional' IVF/ICSI with controlled ovarian hyperstimulation (COH-IVF) and children born after IVF/ICSI in the modified natural cycle (MNC-IVF). A third group was formed by children of couples who had a naturally conceived (NC), singleton pregnancy while on the waiting list for fertility evaluation or treatment during the same time frame. The couples on the waiting list had been subfertile for at least 1 year. The latter control cohort was chosen as we expected that parental characteristics like, for instance, parity and parental age of the subfertile couples would resemble the characteristics of IVF couples. Excluded from the study were twins and children born after cryopreserved or donated oocytes or embryos.

The Groningen ART-cohort is formed to disentangle the separate effects of ovarian hyperstimulation and the *in vitro* procedure. MNC-IVF differs fundamentally from COH-IVF in that the one oocyte that naturally develops to dominance is used. Minimal medication is supplemented to achieve this goal. Potential differences in neurodevelopmental outcome of COH-IVF and MNC-IVF children may, therefore, be attributed to ovarian hyperstimulation and/or the loss of natural selection of the dominant follicle. Likewise, potential neurodevelopmental differences between MNC-IVF and NC-children are attributable to the *in vitro* procedure (figure 1). When results of the Groningen ART-cohort study are interpreted, it should be noted that the medication used in MNC-IVF, although minimal, may cause an overestimation of the effect of the IVF procedure itself and an underestimation of the effect of ovarian stimulation. Since IVF is in general applied following ovarian hyperstimulation, we

Figure 1 - Diagram of the groups and effects studied in the Groningen ART-cohort study



also compared neurodevelopmental outcome of COH-IVF and NC children. The information obtained in this way is valuable for subfertile couples considering IVF treatment (figure 1).

Follow-up of children born after IVF with Preimplantation Genetic Screening

The PGS follow-up study consists of children of couples who participated in a double blind, two-centre, randomised controlled trial on the efficiency of PGS to improve ongoing pregnancy rates in IVF (Mastenbroek *et al.*, 2007). Women participating in the trial had to be between 35 and 41 years old. Exclusion criteria for women were previously failed IVF-cycles and objections against a possible double embryo transfer. Randomisation of women into IVF with or without PGS was performed centrally with minimisation for age (35-37 or 38-41 years) and reproductive technique (IVF or ICSI), with stratification according to study centre (Academic Medical Center, Amsterdam or University Medical Center Groningen) (Mastenbroek *et al.*, 2007). Couples were informed that a follow-up program was part of the PGS-trial and that children born to couples who were included in the trial would be invited for follow-up during the third trimester of pregnancy. For practical reasons, children born after treatment in Groningen were invited to participate in an extensive follow-up program, whereas children born after treatment in Amsterdam were only invited for assessments at the age of two years.

Table 1 - The assessment battery

Assesment age	Neurodevelopmental test used	Outcome measures
2 weeks	GMs	Quality of GMs
3 months	GMs	Quality of GMs
4 months	TINE	Clinical neurological classification
10 months	TINE	Clinical neurological classification
18 months	Hempel	NOS, fluency-score and clinical neurological classification
2 years	Hempel	NOS, fluency-score and clinical neurological classification
	BSID-II	MDI and PDI
	CBCL	Total problem scale, Internalizing scale and Externalizing scale

GMs: General Movements, TINE: Touwen Infant Neurological Examination, Hempel: Hempel's neurological examination for toddlers, NOS: Neurological Optimality Score, BSID-II: Bayley's Scales of Infant Development - second edition, MDI: Mental Developmental Index, PDI: Psychomotor Developemental Index, CBCL: Child Behaviour Check List

An overview of neurodevelopmental tests used in the present thesis

The ages of assessment and neurodevelopmental tests used in the present thesis are presented in table I. Age-specific testing is necessary, since an infant's functional repertoire expands rapidly due to abundant structural changes in the

developing nervous system in the first two years of life. In infancy, the emphasis of the assessments was on quality of neurological functioning, either expressed in the quality of General Movements (GMs) or in the occurrence of Minor Neurological Dysfunction (MND). Complementary, cognitive and behavioural function was assessed at the age of two years.

OUTLINE OF THIS THESIS

The studies in this thesis concern neurodevelopmental outcome of children born following ART. The thesis is divided into 4 parts;

Part I; Literature review
Chapter 2 provides a systematic overview of the literature on neuromotor, cognitive, language and behavioural outcome in children born following IVF or ICSI.

Part II; The Groningen ART-cohort study
In this cohort, we investigated the effect of ovarian hyperstimulation and the in vitro procedure on neurodevelopmental outcome in cohorts of children born following modified natural cycle IVF, controlled ovarian hyperstimulation IVF and children born to subfertile parents.
Chapter 3 describes early neuromotor development, measured by means of the quality of General Movements at 2 weeks and 3 months.
Chapter 4 documents the neurological condition of the children measured with the Touwen Infant Neurological Investigation (TINE) at 4 and 10 months and the Hempel examination at 18 months. Both measures, TINE and Hempel, focus on the presence of minor neurological dysfunction.
Chapter 5 presents the neurological condition of the children measured with the Hempel examination at 2 years.
Chapter 6 describes mental and psychomotor development and behaviour measured with, respectively, the Bayley Scales of Mental Development and the Child Behaviour Check List at the age of 2 years.

Part III; Follow-up of children born after IVF with Preimplantation Genetic Screening
This part addresses neurodevelopmental outcome of children born following IVF with PGS compared to children born after 'conventional' IVF. Neurodevelopmental assessments carried out in the PGS follow-up study are similar to the assessments in the Groningen ART cohort study.

Chapter 7 deals with neurodevelopmental outcome measured at 2 weeks, 3, 4, 10 and 18 months.
Chapter 8 reports mental, psychomotor, neurological and behavioural outcome in 2-year-old children.

Part IV; General discussion, future perspectives and summary
Chapter 9 contains the general discussion and future perspectives.
Chapter 10 summarizes the results of the studies in English and Dutch.

Part I
Literature review

2

Neuromotor, cognitive, language and behavioural outcome in children born following IVF or ICSI - a systematic review

K.J. Middelburg
M.J. Heineman
A.F. Bos
M. Hadders-Algra

Human Reproduction Update 2008; 14: 219-231.

ABSTRACT

Background: The effect of in vitro fertilisation (IVF) or intracytoplasmic sperm injection (ICSI) on the developing human brain is unclear. The objective of this study is to evaluate neurodevelopmental (ND) outcome of children born following these techniques.

Methods: This systematic review includes studies which compare a group of children born following IVF/ICSI to children born after natural conception by assessing outcome in terms of neuromotor development, cognition, speech/language and behaviour. Specific attention is paid to the studies' methodological quality based on study design, attrition, blinding of the assessor, validity of ND tests used, confounders included and group size or power analysis.

Results: Twenty-three out of 59 studies had a good methodological quality including 9 register-based (RB) and 14 controlled studies. RB studies suggested that IVF/ICSI per se does not increase the risk for severe cognitive impairment (i.e. mental retardation) or neuromotor handicaps such as cerebral palsy (CP), the association of IVF/ICSI and CP being brought about by the association of assisted conception with risk factors, like preterm birth. In general, controlled studies of good quality did not report an excess of ND disorders in IVF/ICSI-children. However, the majority of studies followed the children during infancy only, thereby precluding pertinent conclusions on the risk of ND disorders that come to the expression at older ages, such as fine manipulative disability or dyslexia.

Conclusions: A negative effect of assisted conception on the developing human brain is not identified; however, further research of high methodological quality in children beyond pre-school age is needed.

INTRODUCTION

The effect of assisted conception on the developing human brain is still not clear, notwithstanding the fact that in vitro fertilisation (IVF) and intracytoplasmic sperm injection (ICSI) have been introduced more than 25 and 15 years ago, respectively (Steptoe and Edwards, 1978; Palermo *et al.*, 1992). Since then many studies have been conducted on neurodevelopmental outcome (an umbrella term covering neuromotor, cognitive, speech/ language and behavioural outcome) of children born following these techniques but hitherto uncertainties persist. For instance, contradicting results have been reported considering the association between assisted conception and cerebral palsy (CP), which is a neuromotor disorder that is attributed to non-progressive disturbances in the developing brain (Bax *et al.*, 2005). Some studies report an association between assisted conception and CP (Ericson *et al.*, 2002; Lidegaard *et al.*, 2005; Strömberg *et al.*, 2002), whereas others could not demonstrate such association (Pinborg *et al.*, 2004; Källén *et al.*, 2005; Hvidtjørn *et al.*, 2006; Klemetti *et al.*, 2006). The lack of clarity on potential neurodevelopmental risk after IVF/ICSI is worrying for multiple reasons. First, the number of pregnancies obtained by an assisted reproductive technology (ART) is steadily increasing. In Europe, up to 3.9% of national births are infants born after ART (Nyboe Andersen *et al.*, 2007). Moreover, new and more invasive techniques are introduced at a rapid pace and are not always accompanied by extensive follow-up programmes.

Furthermore, there might be reasons to suppose that IVF/ICSI is associated with an increase in neurodevelopmental problems. Early development of the human nervous system is a complex and neatly orchestrated process which can be affected easily by external influences (De Graaf-Peters and Hadders-Algra, 2006). It has already been established that perinatal outcome of singletons born after assisted conception is worse than that of naturally conceived singletons. Artificially conceived singleton pregnancies end significantly more often preterm and with low birth weight (Schieve *et al.*, 2002; Helmerhorst *et al.*, 2004; Jackson *et al.*, 2004) and perinatal mortality and neonatal intensive care admission are increased (Helmerhorst *et al.*, 2004; Jackson *et al.*, 2004).

The lack of consensus about neurodevelopmental risk after IVF/ICSI largely stems from the fact that the results of the various follow-up studies often are difficult to interpret because of methodological shortcomings. Previous reviews mentioned methodological drawbacks but lacked a thorough methodological evaluation (Van Balen, 1998; Buitendijk, 1999; Tarlatzis and Grimbizis, 1999; Olivennes *et al.*, 2002; Ludwig and Diedrich, 2002; Ludwig *et al.*, 2006; Sutcliffe and Ludwig, 2007). Others focused on subgroups of children, e.g. twins (Pinborg,

2005), children born following ICSI (Van Steirteghem *et al*., 2002; Leslie, 2004), or children born following cryopreservation of embryos (Sutcliffe, 2000; Wennerholm, 2000), or took into account only a part of neurodevelopmental outcome, such as psychosocial well being (Hahn, 2001; Colpin, 2002; Golombok and MacCallum, 2003; Gibson and McMahon, 2004). Therefore, the aim of the present review is to evaluate in a systematic manner studies on neurodevelopmental outcome of children born following IVF or ICSI compared to naturally conceived children. We restricted ourselves to the techniques of IVF and ICSI as the character of these procedures is invasive and a substantial number of follow-up studies have been reported. For the still more invasive techniques, such as preimplantation genetic screening and in vitro maturation, follow-up information is almost completely lacking. We first evaluated the methodological quality of the studies in a strict and standardised way. The identified studies of good methodological quality are summarized and the results are presented and discussed in an age-specific manner.

METHODOLOGY

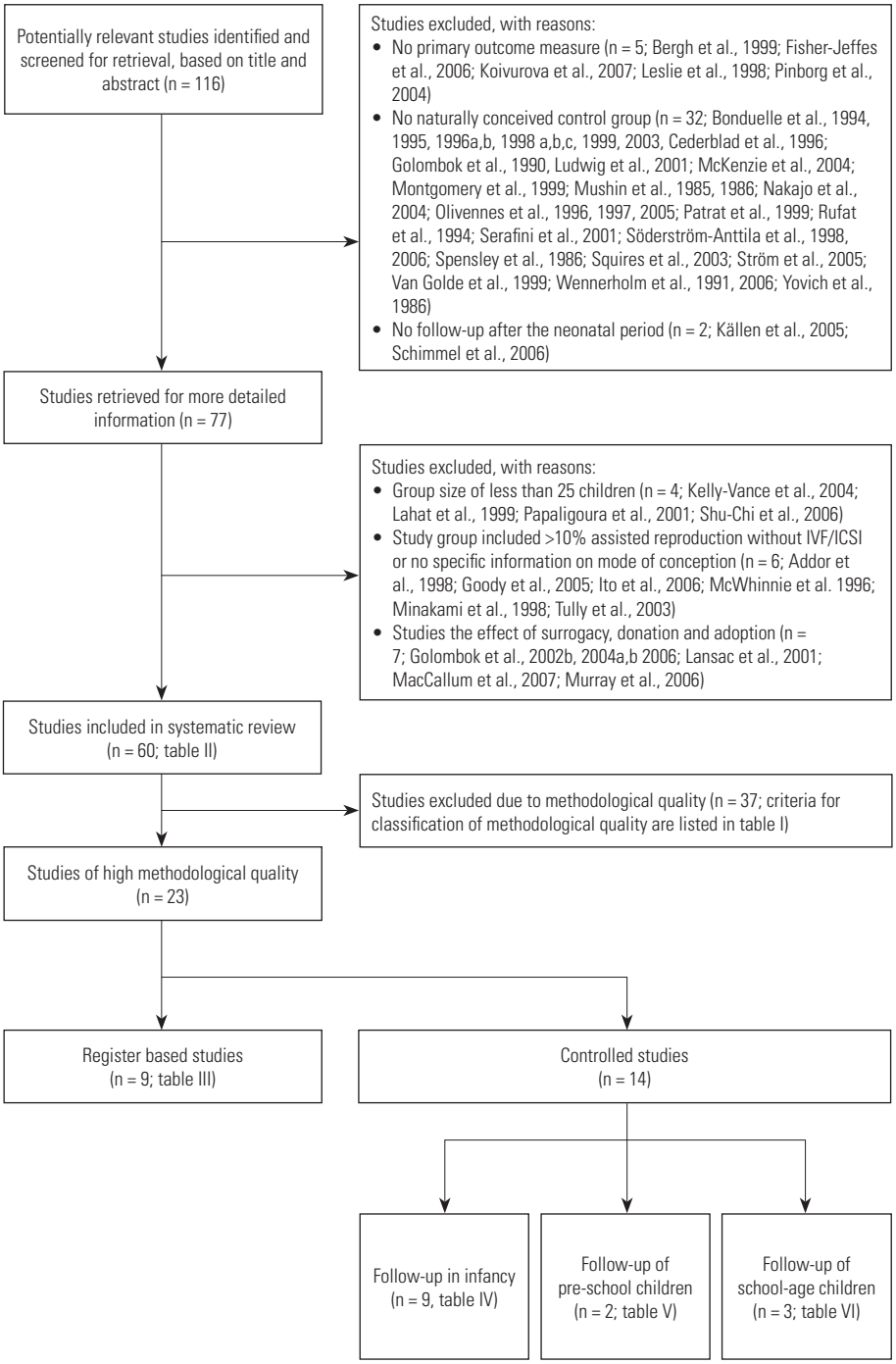
Literature search

An extensive literature search was performed for relevant studies on neurodevelopmental outcome of children born following IVF or ICSI. We searched for articles published between 1978 and 5 December 2007 in Medline, Embase (since its first coverage year: 1989), PsycINFO and the Cochrane library. A computer based search strategy with multiple combinations of terms was entered into the databases. This search strategy consisted of all combinations of (i) IVF OR ICSI OR reproductive techniques, assisted OR fertilisation in vitro AND (ii) child development OR abnormalities OR morbidity OR psychomotor performance OR motor skills OR intelligence OR child psychology OR child behaviour OR developmental disabilities OR nervous system diseases OR CP AND (iii) infant(s) OR child(ren) OR adolescent(s) OR twins OR triplets. Note that the terms were adapted to terminology used in the various databases. In addition, the reference lists of all identified studies and review articles were reviewed for additional articles.

Inclusion and exclusion criteria

We searched for all studies which assessed neurodevelopmental outcome, i.e. neuromotor development, cognition, speech/language and behaviour as a primary outcome measure in IVF or ICSI children and a naturally conceived comparison group.

Figure 1 - Flow-diagram of included studies



Excluded from the study were studies (i) which did not include a naturally conceived control group, (ii) with a study group size of less than 25 children, (iii) in which the follow-up did not extend beyond the neonatal period, (iv) in which the study group included more than 10% of children born following ovulation induction only (without IVF or ICSI), (v) which compared outcome of children born following IVF or ICSI to children born from donor gametes, adopted children or children born from surrogate mothers and (vi) not published in English. The decision to exclude studies with a group size of less than 25 children was based on preliminary results which revealed that a small group size virtually always was associated with a poor methodological quality. Family studies were only included when a substantial part of the study was devoted to the child's neurodevelopmental outcome.

Identification

The search strategy yielded 1131 publications in Medline and Embase, 213 in PsycINFO and 181 in the Cochrane library. These were supplemented with articles found in reference lists. On the basis of abstract and title, 115 potentially relevant articles were identified and screened for retrieval. Figure 1 shows a flow-diagram of the in and excluded studies, with reasons for exclusion. The manuscripts of the studies included in the systematic review were read in full by two independent reviewers (M.H-A. and K.J.M.). Study characteristics, data qualifying methodology and data on outcome were extracted and discussed until consensus was met.

Methodological hierarchy

Studies evaluating outcome after IVF/ICSI do not allow for a randomised clinical trial (Buck Louis *et al.*, 2005). The best option for a clinical trial is the design which evaluates outcome in prospective cohorts of consecutively born IVF/ICSI children and naturally conceived controls, both recruited pre- or perinatally. Next best approaches are studies in which IVF/ICSI children are studied prospectively as a cohort, but the naturally conceived controls are matched retrospectively at nursery- or school age. In order to enhance the ability to fine grade the quality of IVF/ICSI studies, we therefore made the differentiation between prospective- and retrospective-cohort studies (respectively, PC and RC) according to the enrollment of the control children. Studies, which included IVF/ICSI children whose selection was not clearly defined, i.e. studies in which it was not clear whether the children studied represented the entire population of a region or a centre, were classified as retrospective cohort. Studies which included children with a disorder or disease (e.g. CP) and evaluated the mode of conception of children in the diseased and non-diseased groups were classified as case–control (CC).

The effect of IVF/ICSI has not only been studied in clinical trials, but also in studies

Table I - Criteria for classification of methodological quality of controlled studies.

Criteria for:	Assessed criteria:	Qualification:
Confounders accounted for	Either by matching in study protocol including report of success of matching, or by means of multivariate statistical analysis: i) plurality ii) gestational age/ low birth weight iii) parity iv) maternal age v) parental education or social-economical class vi) testing age	+++ = ≥ 5 parameters. ++ = 4 parameters. + = 3 parameters. - = ≤ 2 parameters.
Internal validity	i) prospective cohort study or register based study ii) postnatal attrition ≤ 10% iii) blindness of the assessor iv) good validity of the neurodevelopmental test (either generally acknowledged or documented in paper) v) ≥ 4 confounders taken into account	+++ = 4 or 5 criteria fulfilled ++ = 3 criteria fulfilled + = 2 criteria fulfilled - = 0-1 criteria fulfilled
Power of study to assess neurodevelop-mental (ND) outcome	i) power analysis ii) sample size	+ = power analysis provided, or samples evaluated larger than 2,500 individuals - = power analysis not provided and samples evaluated smaller than 2,500 individuals.
External validity	i) qualification of internal validity ii) qualification of power of study to assess neurodevelopmental outcome	+++ = internal validity +++ and power + ++ = (internal validity ++ and power +) or (internal validity +++ and power – with sample sizes ≥ 50 subjects) + = (internal validity + and power +) or (internal validity +++ and power – with samples sizes < 50 subjects) or (internal validity ++ and power – with sample sizes ≥ 50 subjects) - = not fulfilling the criteria for +++/ ++/ +

based on nation-wide registers. These register-based (RB) studies are in particular valuable to detect disorders of low incidence, like CP. Hierarchically, RB studies were considered as having the same level of evidence as the PC studies.

Attrition is an important problem in follow-up studies. We classified studies according to their degree of post-natal attrition. This means that attrition due to perinatal mortality is not taken into account. A low rate of attrition indicates that the chance of selection bias will be kept to a minimum. Blind evaluation of outcome is another important criterion for validity. We only considered a study blind if this was explicitly mentioned in the paper or if data were collected for a purpose other than

the evaluation of the effect of IVF/ICSI treatment. A case in point is data collected in Child Welfare Clinics or in habilitation centres.

Neurodevelopmental outcome is affected by multiple factors. Studies dealing with the effect of IVF/ICSI should take into account these confounding factors, either by matching study- and control groups or by including the confounders in multivariate statistics evaluating outcome. We evaluated whether studies took the following confounding variables into account: plurality, gestational age or low birthweight, parity, maternal age, parental education, parental profession or other indicators of social–economical class and the age at which the child had been tested. We did not list gender amongst the confounders, as we were interested in particular whether boys or girls might have a specific vulnerability for sequelae after ART. It is well known that the prevalence of neurodevelopmental disorders is higher in boys than in girls (Hadders-Algra *et al.*, 1988a,b). Therefore, we chose to explicitly report the results if studies did stratify for gender.

Quality assessment was performed by the two independent reviewers and took into account internal validity (quality of the study) and external validity (generalisability of the study; Moher *et al.*, 1999). Internal validity was based on (i) design of the study, (ii) attrition, (iii) blinding at evaluation, (iv) validity of the neurodevelopmental tests applied in the study and (v) the degree to which confounders were taken into account. External validity was based on internal validity, the size of the samples studied and/or the presence of a power calculation. When results of a study had been reported in two different papers, only the paper which described the study most extensively was taken into account for the review. Criteria for classification of methodological quality of controlled studies, i.e. internal and external validity, are listed in Table I.

Presentation of the results

We first report the methodological quality of the retrieved studies (table II). Thereafter, we present the data of the studies with a good methodological quality (≥ ++ external validity). A differentiation was made between RB (table III) and controlled studies (Tables IV, V, VI). For the controlled studies, infants, pre-school children and school-age children are reported separately as neurological dysfunction in these age groups is expressed differently. This is for instance reflected by the existence of specific neurological assessments for infants, pre-schoolers and school-age children (Hadders-Algra, 2005). Owing to the plasticity of the brain dysfunctions that exist at a young age may disappear when the child grows older. But the reverse may also occur: with increasing age increasingly complex brain functions become functionally expressed, which may be accompanied by the emergence of dysfunctions in the novel functions. A case in point is the development of dyslexia at school age (Hadders-Algra, 2005). In the different age-groups, neuromotor development (including

neuromotor handicaps), cognition, language and behaviour are reported separately. Ages of the infants at evaluation, neurodevelopmental tests used and study outcome were recorded. Owing to the heterogeneity in tests used and child age a meta-analysis on the effect of IVF/ICSI could not be performed.

RESULTS

Methodological quality of the studies

Fifty-nine studies fulfilled the selection criteria, including 9 RB studies and 50 controlled studies. Study characteristics are summarized in Table II. Sample size of the study groups ranged from 26 to 16 280 children. There was a considerable amount of overlap of children between studies. We decided to report the studies separately as they in general used different developmental outcome parameters. Most studies included children born following conventional IVF or a mixture of IVF and ICSI (denoted by S; 33 out of 59 studies, i.e. 56%). Twenty-four studies reported results of children born following ICSI only (S1; 41%) and two studies reported on children born following cryopreservation as embryos (S2; 3%).

Twenty-three of the controlled studies used a prospective study design for both study and control group. In these studies, the control group was formed as a cohort either prenatally, neonatally or from a hospital delivery register. Twenty-seven cohort studies were performed retrospectively. This always meant that at least a part of the control group had been recruited retrospectively, e.g. via nurseries or schools. A CC design was not used in the controlled studies. Information on attrition was available in 97% of the studies; it varied from 0 to 88%.

Thirty-three of the identified studies had good (++ or +++) internal validity and 23 studies had good external validity. Only outcome of studies with good external validity are reported in detail, separately for RB studies (9; 100%) and controlled studies (14; 28%).

RB studies

In some Scandinavian countries, national registers of fertility treatment have been linked to national registers of hospital diagnoses or registers of psychiatric disorders to calculate the odds for adverse outcome in children born following fertility treatment. Nine of these RB studies focused on neurodevelopmental outcome (table III). Three studies included only children aged at least 1½ or 2 years to ensure accurate neurological diagnosis (Strömberg *et al.*, 2002; Pinborg *et al.*, 2004; Klemetti *et al.*, 2006). With respect to risk for CP, this is very appropriate as

Table II - Study characteristics.

Author and publication year	Studie group* (n)	Control group (n)	Plurality †	Sample origin	Control origin
Agarwal et al. 2005	S1 = 76	261	sin, tw, tri	clinical	hospital delivery register
Audiens et al. 1995	S = 50	50	sin, tw	clinical	referral polysomniography
Barnes et al. 2004	S = 301, S1= 345	310	sin	multicenter	nursery + birth register
Belva et al. 2007	S1 = 150	147	sin	clinical	school-age peers
Bonduelle et al. 2004	S1 = 300	266	sin	multicenter	school-age peers
Bonduelle et al. 2005	S = 437, S1 = 540	538	sin	multicenter	nursery + birth register
Bowen et al. 1998	S = 84, S1 = 89	80	sin, tw	clinical	prenatal peers
Brandes et al. 1992	S = 116	116	sin, tw, tri	clinical	hospital delivery register
Colpin et al. 1995	S = 31	31	sin	clinical	hospital delivery register
Colpin & Soenen 2002	S = 27	23	sin	clinical	hospital delivery register
D'Souza et al. 1997	S = 278	278 sin	all	clinical	hospital delivery register
Ericson et al. 2002	S = 9056	1417166	all	national	national register
Gershoni-Baruch et al. 1991	SA= 33 (HFTVS), SB = 45	33 + 45	sin, tw	clinical	infant peers
Gibson et al. 1998	S = 65	62	sin.	clinical	prenatal peers
Gibson et al. 2000	S = 65	61	sin	clinical	prenatal peers
Golombok et al. 1995	S = 41	43	sin	multicenter	hospital delivery register
Golombok et al. 1996	S = 116	120	sin	multicenter	hospital delivery register + school-age peers
Golombok et al. 2001	S = 34	38	sin	multicenter	hospital delivery register
Golombok et al. 2002	S = 102	102	sin	multicenter	hospital delivery register + school-age peers
Hahn & DiPietro 2001	S = 54	59	sin	multicenter	school-age peers
Hvidtjørn et al. 2006	S = 9255	394713	sin, tw	national	national register
Källén et al. 2005	S = 16280	population	all	national	national register
Klemetti et al. 2006	S = 4559	26877	sin, tw, tri	national	national register
Knoester et al. 2007a	S = 81, S1 = 87	85	sin	clinical	school-age peers
Knoester et al. 2007b	S = 81, S1 = 87	85	sin	clinical	school-age peers
Knoester et al. 2007c	S = 86, S1 = 83	85	sin	clinical	school-age peers
Koivurova et al. 2003	SA = 299 SB = 250	CA = 558 CB = 380	A = all B= sin, tw	multicenter	birth register
La Sala et al. 2004	S1 = 50	51	sin, tw	clinical	peers via paediatrician
Leslie et al. 2003	S = 80, S1 = 97	110	sin, tw.	clinical	nursery + prenatal peers

Study design ‡	Postnatal attrition FU§	Blind ‖	Validity of ND tests used	Conf. acc. for**	Internal validity**	Power ND test**	External validity**
PC	3-10%	yes	++	++	+++	+	+++
RC	0%	?	++	-	+	-	-
RC	4-66%	no	++	++	+	-	-
RC	39-62%	no	++	++	+	-	-
RC	12,5-70%	partially	+	-	-	-	-
RC	4-75%	partially	+	++	-	-	-
PC	2–3%	no	++	+++	+++	-	++
PC	0-13%	yes	++	+++	+++	-	++
PC	11%	no	+	++	+	-	-
PC	13-25% P 34-52% T	P: no T: yes	++	++	++ (P) +++ (T)	-	+ (P) + (T)
PC	?	?	++	-	+	-	-
RB	<1%	yes	++	+	+++	+	+++
RC	~0%	?	++	+	+	-	-
PC	2 -7%	?	++	+++	+++	-	++
PC	3-7%	yes	++	+++	+++	-	++
PC	≥ 5-38%	P: no T: yes	++ Rutter + Interv. - m SAT + PC&SA	++	++	-	+
RC	≥ 33%	P: no T: yes	++ Rutter + PC&SA	++	+	-	-
PC	≥11-17%	P: no T: yes C: ?	+ Psych. Int. ++ SDQ + SAICA	++	++	-	+
RC	≥ 12-15%	P: no T: yes C: ?	++ SDQ + CAFÉ	++	+	-	-
RC	≥24%	P: no T: yes	++ ECBI ++ PBCL + SESBI - TRRLS	++	+	-	-
RB	< 1%	yes	++	+++	+++	+	+++
RB	<1%	yes	++	++	+++	+	+++
RB	<1%	yes	++	+	+++	+	+++
RC	21- 33%	yes	++	+++	++	+	++
RC	21-27%	no	++	+	-	-	-
RC	21- 33%	yes	++	+++	++	+	++
PC	10-15%	yes	+	-	+	+	+
RC	40-58%	yes	++	++	++	-	+
RC	5-88%	yes	++	++	++	+	++

Table II - *Continued*

Author and publication year	Studie group* (n)	Control group (n)	Plurality †	Sample origin	Control origin
Leunens et al. 2006	S1 = 151	153	sin	clinical	school-age peers
Leunens et al. 2007	S1 = 109	90	sin	clinical	school-age peers
Levy-Shiff et al. 1998	S = 51	51	?	multicenter	school-age peers
Lidegaard et al. 2005	S = 6052	442349	sin	national	national register
Maimburg & Væth 2007	S = 461	461	all	national	national register
McMahon et al. 1997	S = 65	62	sin	clinical	prenatal peers
McMahon & Gibson 2002	S = 70	63	sin	clinical	prenatal peers
Morin et al. 1989	S = 83	93	sin, tw, tri	clinical	hospital delivery register
Neri et al. 2004	S1 = 101	57	sin	clinical	school-age peers
Papaligoura et al. 2004	S = 26, S1 = 34	29	sin, tw	clinical	peers via paediatrician
Pinborg et al. 2003	SA = 634 sin SB = 472 tw	1132 tw	sin, tw	national	birth register twins
Pinborg et al. 2004	SA = 5130 sin SB = 3393 tw	10239 tw	sin, tw	national	national register twins
Place & Englert 2003	S = 52, S1 = 66	59	sin	clinical	hospital delivery register
Ponjaert-Kristoffersen et al. 2004	S1 = 300	260	sin	multicenter	nursery + birth register
Ponjaert-Kristoffersen et al. 2005	S = 424, S1= 511	488	sin	multicenter	nursery + birth register
Raoul-Duval et al. 1993, 1994	S = 33	33	sin	clinical	neonatal peers
Ron-El et al. 1994	S = 32	32	sin	clinical	neonatal peers
Sanchez-Albisua et al. 2007	S1 = 34	39	sin	clinical	hospital delivery register
Saunders et al. 1996	S = 289	146	sin, tw	multicenter	hospital delivery register
Strömberg et al. 2002	S = 5680	15397	all	national	national register
Sutcliffe et al. 1995a,b	S2 = 91 (cryo)	83	sin, tw, tri	clinical	nursery peers
Sutcliffe et al. 1999	S1 = 123	123	sin	clinical	nursery peers
Sutcliffe et al. 2001	S1 = 208	221	sin	multicenter	nursery peers
Sutcliffe et al. 2003	S1 = 264	260	sin	multicenter	nursery peers
Sutcliffe et al. 2004	S1 = 140	101	sin	multicenter	nursery peers
Sutcliffe et al. 2005	S = 425 S1 = 525	523	sin	multicenter	nursery + birth register
Sun et al. 2007	S = 1958	50396	sin	national	national register
Sydsjö et al. 2002	S = 121	110	sin, tw	clinical	prenatal peers
Van Balen 1996	S = 45	35	sin	clinical	prenatal peers
Wennerholm et al. 1998	S = 255 (fresh), S2 = 255 (cryo)	252	sin, tw.	clinical	hospital delivery register

*S = studygroup; children born following IVF or a combination of IVF and ICSI, S1 = children born following ICSI, S2 = children born following cryopreservation as embryos. HFTVS = High-frequency transvaginal ultrasonography.

† sin = singletons, tw = twins, tri = triplets, all = children of all pluralities.

‡ PC = prospective cohort study, RC = retrospective cohort study, CC = case control study, RB = register based study.

§ Postnatal attrition follow-up, P = parent, T = teacher.

|| P = parent, T = teacher , C = child.

Study design ‡	Postnatal attrition FU§	Blind ‖	Validity of ND tests used	Conf. acc. for**	Internal validity**	Power ND test**	External validity**
RC	39-62%	no	++	+++	+	+ WISC - mABC	+ (WISC) - (mABC)
RC	41-56%	No	++	++	+	-	-
RC	0-8%	yes	++	++	+++	-	++
RB	<1%	yes	++	-	+++	+	+++
RB	0%	yes	++	++	+++	-	++
PC	<5%	yes SFP no Q	++ SFP + Q	+++	+++ SFP ++ Q	-	++ SFP + Q
PC	<10%	yes SSP no Q	++ SSP + Q	++	+++ SSP ++ Q	-	++ SSP + Q
PC	11-25%	yes	++	++	+++	-	++
RC	?	?	++	-	-	-	-
RC	0 -7%	yes	++	+++	+++	-	+
PC	11- 23%	no	+	+++	+	-	-
RB	<1%	yes	++	++	+++	+	+++
PC	30-60%	no	++	+++	++	-	+
RC	12-70%	?	++	++	+	-	-
RC	4-66%	partially	++	++	+	-	-
PC	6-67%	yes	++	++	+++	-	+
PC	3-19%	yes	++	+++	+++	-	+
PC	~50-72%	?	+	++	+	-	-
PC	≥27-88%	no	++	+	+	-	-
RB	< 1%	yes	++	++	+++	+	+++
RC	~0%	no	++	-	+	-	-
RC	10%	?	++	++	++	+	++
RC	1-10%	no	++	+++	++	+	++
RC	10-15%	no	++	+	-	+	-
RC	7-10%	no	++	-	+	-	-
RC	0-75%	yes	++	-	+	-	-
RB	~10%	yes	++	+++	+++	-	++
PC	2%	no	+	++	++	-	+
PC	31-65%	no	+	+++	+	-	-
PC	0-2%	yes	+	-	++	-	+

¶ Neurodevelopmental tests used: Rutter = Rutter's behaviour scale, Interv. = Interview, mSAT = modified Seperation Anxiety Test, PC&SA = Pictural Scale for Perceived Competence and Social Acceptance for Young Children, SDQ = Strengths and difficulties questionnaire, SAICA = Social Adjustment Inventory for Children and Adolescents, CAFÉ = Child and Adolescent functioning and Environment Schedule, ECBI = Eyberg Child behaviour Inventory, PBCL = Pre-school Behaviour Checklist, SESBI = Sutter-Eyberg Student Behaviour Inventory, TRRLS = Teachers report on response to limit setting, SFP = Still-face Procedure, Q = Questionnaire, SSP = Strange Situation Procedure

** for criteria see table I

Table III - Register based studies on neurodevelopmental outcome of children born following IVF/ICSI.

Study	Age	Methodological quality*		Variables controlled for †
		Int*	Ext*	
Pinborg et al. 2004	2-7 years	+++	+++	Maternal age, ART procedure, plurality, LBW, preterm birth, gender, year of birth.
Strömberg et al. 2002	1½ - 14 years	+++	+++	Maternal age, ART procedure, plurality, LBW, preterm birth, gender, year of birth, birth hospital
Klemetti et al. 2006	2- 4 years	+++	+++	Maternal profession, plurality
Hvidtjørn et al. 2006	1-7 years	+++	+++	Maternal age, maternal educational level, parity, plurality, preterm birth, SGA status, gender.
Källén et al. 2005	1-20 years	+++	+++	Maternal age, smoking habit, parity, years of unwanted childlessness, preterm birth, year of birth.
Ericson et al. 2002	1-14 years	+++	+++	Maternal age, smoking habit, parity, year of birth.
Lidegaard et al. 2005	1-7 years	+++	+++	Plurality
Sun et al. 2007	0-6 years	+++	++	Maternal age, paternal age, maternal social status, smoking habit, BMI, parental epilepsy, parity, plurality, preterm birth, year of birth.
Maimburg & Væth 2007	0-11 years	+++	++	Maternal age, maternal origin, parity, plurality, preterm birth, birth weight, birth defects.

* Methodological quality: see table II. Int. = Internal validity, Ext. = External validity

† ART = assisted reproductive technology, LBW = low birth weight, SGA = small for gestational age, BMI = body mass index.

	Neurodevelopmental outcome ‡	
Register used	**Outcome after first correction for confounders**	**Outcome after final correction for confounders**
National patient/ psychiatric register (ICD-10)	CP: $S_{singletons} = S_{twins} = C_{twins}$ Mental retardation: $S_{singletons} = S_{twins} = C_{twins}$	CP: $S_{singletons} = S_{twins} = C_{twins}$
National register of diagnosis at childhood disability centre (ICD-10)	CP: S > C (singletons), S = C (twins) Mental retardation: S = C Developmental delay: S = C Behavioural disorder: S = C	CP: S = C (singletons), S > C (all pluralities) Developmental delay: S > C (all pluralities) Developmental delay: S = C (singletons)
National hospital discharge register (ICD-10)	CP: S > C (all pluralities) Behavioural disorder: S = C (all pluralities) Epilepsy: S > C (all pluralities)	CP: S = C (singletons), S = C (twins) Behavioural disorder: S = C (singletons), S = C (twins) Epilepsy: S = C (singletons), S = C (twins)
National hospital discharge register (ICD-10)	CP: S > C	CP: S = C
National hospital discharge register (ICD-9/10)	CP: S > C Mental retardation: S = C Behavioural problems: S > C Epilepsy: S > C Convulsion: S > C	CP: S = C Mental retardation: S = C Behavioural problems: S = C Epilepsy: S = C Convulsion: S > C
National hospital discharge register (ICD-9/10)	Not reported.	CP: S > C Mental retardation: S = C Developmental disturbances: S = C Epilepsy: S > C
National patient/ psychiatric register (ICD-10)	CP: S > C Mental retardation: S = C Behavioural disturbance: S = C Sleeping disturbance: S > C Speech/ language retardation: S = C Motor retardation: S = C	n.a.
National hospital discharge register (ICD-10)	Epilepsy: S > C Febrile seizures: S = C	Epilepsy: S = C Febrile seizures: S = C
National patient/ psychiatric register (ICD-8/10)	Infantile Autism: S < C	Infantile Autism: S < C

‡ ICD-9/10 = International Classification of Diseases, 9th/10th revision, CP = Cerebral Palsy, S = study group, C = control group, S = C: no statistically significant differences between study and control group. S > C: significantly more problems in study than in control group, S < C, significantly less problems in study than in control group, n.a.= not applicable

the diagnosis CP cannot be established prior to the age of 1½–2 years of age (Bax *et al.*, 2006). Strömberg *et al.* (2002) identified IVF as an independent risk factor for development of CP, but the effect disappeared when only singletons were taken into account. Pinborg *et al.* (2004) reported that twins born following IVF/ICSI had a similar risk of CP as naturally conceived twins and singletons born following IVF/ICSI. Klemetti *et al.* (2006) reported an increased risk of CP for children born after IVF/ICSI when all pluralities were taken into account, but not for singletons only. Studies that included also younger children (Ericson *et al.*, 2002; Källén *et al.*, 2005; Lidegaard *et al.*, 2005; Hvidtjørn *et al.*, 2006) showed that IVF/ICSI might be associated with a higher prevalence of CP, but when the results had been adjusted for important confounders such as preterm birth and plurality the increased risk usually disappeared. Interestingly, epilepsy and/or the occurrence of convulsions remained associated with IVF/ICSI in some studies even after correction for confounders (Ericson *et al.*, 2002; Källén *et al.*, 2005).

The association of IVF/ICSI and mental retardation was investigated in five papers. These studies reported a consistent absence of a relationship between IVF/ICSI and mental retardation. Speech/language retardation was only investigated by Lidegaard *et al.* (2005); no differences between IVF/ICSI children and naturally conceived children were demonstrated. Four studies reported on behaviour, three of them found no differences and one—the CC study of Maimburg and Væth (2007)—surprisingly reported that children born after assisted conception had a lower risk of developing infantile autism.

None of the RB studies reported different outcomes after IVF/ICSI for boys and girls.

Controlled studies with follow-up in infancy

Nine studies with good external validity (41%) investigated neurodevelopmental outcome of infants born following assisted conception (table IV). Five evaluated children born following routine IVF (or a combination of IVF and ICSI) and four studies focused on infants conceived by ICSI. The age of the infants at the time of assessment varied from 4 months to two and a half years. In four studies, the same group of infants was examined (McMahon *et al.*, 1997; Gibson *et al.*, 1998, 2000; McMahon and Gibson, 2002).

Neuromotor development was assessed in six studies either with the help of the Bayley Scales of Infant Development (second edition: BSID-II; Bayley, 1993) or with the Griffiths mental development scales (Griffiths, 1996). Four studies found no differences between infants born following assisted and natural conception. One study reported that children born after ICSI had significantly worse eye–hand coordination than naturally conceived infants (Sutcliffe *et al.*, 1999). Another study

Table IV - Controlled studies with neurodevelopmental outcome in infancy (1 month- 2½ year) and good external validity.

Study	Age	Methodological quality*		Neuromotor development/handicaps		Cognition		Speech/language		Behaviour	
		Int.	Ext.	test used†	outcome‡	test used†	outcome‡	test used†	outcome‡	test used†	outcome‡
Agarwal *et al.* (2005)	2 years	+++	+++	Bayley	S_1=C	Bayley	S_1=C	-		VABS	S_1=C
Gibson *et al.* (2000)	1 years	+++	++	–		–		–		SSP EAS	S=C S=C
Gibson *et al.* (1998)	1 years	+++	++	Bayley	S=C	Bayley	S=C	REEL-2; expressive: receptive:	 S=C S<C	Bayley STST BCL VABS; soc. domain	S=C S<C S<C S=C
McMahon *et al.* (1997)	4 months	+++	++	–		–		–		SFP STSI NPI	S<C S<C S=C
McMahon and Gibson (2002)	1 years	+++	++	–		–		–		SFP SSP	S<C S=C
Morin *et al.* (1989)	1–2½ years	+++	++	Bayley Neuro-paed. exam.	S>C S=C	Bayley	S=C	–		Bayley; vocalization and energy levels:	S=C S>C
Bowen *et al.* (1998)	1 years	+++	++	Bayley	$S=S_1=C$	Bayley	S_1<(S=C) ♀: $S=S_1=C$ ♂: S_1<(S=C)	-		-	
Sutcliffe *et al.* (1999)	1–2 years	++	++	Griffiths; eye–hand coord.:	S_1=C S_1<C	Griffiths	S_1=C	Griffiths	S_1=C	–	
Sutcliffe *et al.* (2001)	1–2 years	++	++	Griffiths Neuro-paed. exam.	S_1=C S_1 =C	Griffiths	S_1=C	Griffiths	S_1=C	-	

* Methodological quality: see table II. Int. = Internal validity, Ext. = External validity

† Neurodevelopmental tests used: Neuro-paed. exam = general neuro-paediatric examination, Bayley = Bayley's Scale of Infant Development (second edition; BSID-II), Griffiths = Griffiths scales of mental development, REEL-2 = Receptive-Expressive Emergent Language Test, 2nd edition, VABS = Vineland Adaptive Behaviour Scale, SSP = Strange Situation Procedure, EAS = Emotional Availability Scales, STST = Short Temperament Scale for Toddlers, BCL = Behaviour Checklist, SFP = Still-face procedure (mother-child interaction), STSI = Short Temperament Scale for Infants, NPI = Neonatal Perception Inventory.

‡ S = study group; infants born following IVF or IVF/ICSI, S1 = study group; infants born following ICSI, C = control group, S = C: no statistically significant differences between study and control group. S > C = Study group performs significantly better than control group, S < C Study group performs significantly worse than control group

indicated that children born after IVF had a better psychomotor development than naturally conceived controls (Morin *et al.*, 1989). Two studies provided information on neurological handicap. They found similar rates of handicap in children born following IVF/ICSI and controls.

Cognition was assessed in six studies: four used the Bayley scales and two the Griffiths scales. Five out of the six studies reported no differences between the study and the control group in mental development. Only the study of Bowen *et al.* (1998) reported significantly lower mental scores in 1-year-old infants born after ICSI than in age matched infants born after IVF and naturally conceived infants. Stratification for gender revealed that lower mental development index scores were only found in boys but not in girls (table IV).

Speech and language were tested in three studies. The results were inconsistent: two studies showed no differences between study and control group and one study found lower scores on receptive language development in infants born following IVF, but no differences in expressive language skills (Gibson *et al.*, 1998).

Behaviour was assessed in six studies. It is good to realize that many different inventories are available to test behaviour. Most objective evaluation procedures are observational measures in which scores are based on observed behaviour during testing. Four of the six studies evaluated behaviour of the same group of infants, be it at different ages (McMahon *et al.*, 1997; Gibson *et al.*, 1998, 2000; McMahon and Gibson, 2002). These four studies in general showed little difference between study and control infants, but on some inventories the IVF-infants were rated as having a more difficult temperament than naturally conceived infants (McMahon *et al.*, 1997; Gibson *et al.*, 1998; McMahon and Gibson, 2002). Two of the six studies reported no differences in behaviour and one of the six studies reported higher energy levels and vocalization in infants born after IVF (Morin *et al.*, 1989).

In summary, the controlled studies with follow-up in infancy do not indicate that neuromotor or cognitive development, including handicapping neurological conditions, of infants born after IVF/ICSI differs from that of non-IVF/ICSI controls. Nor do these studies indicate that infants born following IVF/ICSI show a higher prevalence of language- or behaviour problems than naturally conceived controls.

Controlled studies with follow-up of pre-school children

Only two studies on neurodevelopmental outcome at pre-school age had a good external validity (Brandes *et al.*, 1992; Leslie *et al.*, 2003; Table V). The other 13 studies with follow-up at pre-school age did not fulfil the criteria for good external validity, mainly because of a retrospective design, high attrition and non-blinded

Table V - Controlled studies with neurodevelopmental outcome in pre-school children (3 - 5½ years) and good external validity

Study	Age	Methodological quality*		Neuromotor development/ handicaps		Cognition		Speech/language		Behaviour	
		Int.	Ext.	test used†	outcome‡	test used†	outcome‡	test used†	outcome‡	Test used†	Outcome‡
Brandes *et al.* 1992	1–4 y	+++	++	Neuro-paed. exam.	S=C	Bayley Stanford-Binet	S=C S=C	–	–	–	–
Leslie *et al.* 2003	5 y	++	++	–	–	WPPSI—R	$S=S_1=C$	–	–	–	–

* Methodological quality: see table II. Int. = Internal validity, Ext. = External validity.

† Neurodevelopmental tests used: Neuro-paed. exam = general neuro-paediatric examination, Bayley = Bayley's Scale of Infant Development (second edition; BSID-II), Stanford-Binet = Stanford-Binet intelligence scale, WPPSI-R = Wechsler Preschool and Primary Scales of Intelligence- Revised

‡ S = study group; infants born following IVF or IVF/ICSI, S1 = study group; infants born following ICSI, C = control group, S = C: no statistically significant differences between study and control group. S > C = Study group performs significantly better than control group, S < C Study group performs significantly worse than control group.

testing. The study of Leslie *et al.* (2003) was a follow-up of the children described by Bowen *et al.* (1998). It revealed that the previously reported difference in cognitive development at 18 months between children born after ICSI compared to those born after IVF or after natural conception had disappeared at 5 years. The other study with follow-up at pre-school age did not find a difference in neurological handicap and cognition between children born after IVF and controls (Brandes *et al.*, 1992). Neither of the studies stratified for gender.

Controlled studies with follow-up of school-age children

Three studies with good external validity reported on neurodevelopmental outcome of school-age children (Levy-Shiff *et al.*, 1998; Knoester *et al.*, 2007a, 2008; Table VI). Major reasons for insufficient external validity of the other 10 studies were high attrition—a generally recognised problem in long term follow-up studies—and the use of non-validated outcome measures. The study of Levy-Shiff *et al.* (1998) reported that children born after IVF/ICSI in general do not differ from their naturally conceived peers. Nevertheless, the data indicated that children born after IVF had more socioemotional problems, aggression, anxiety and depression than

Table VI - Controlled studies with neurodevelopmental outcome in school-age children (> 6 years) and good external validity

Study	Age	Methodological quality*		Neuromotor development/handicaps		Cognition		Speech/language		Behaviour	
		Int.	Ext.	Test used†	Outcome‡	Test used†	Outcome‡	Test used†	Outcome‡	Test used†	Outcome‡
Levy-Shiff *et al.* (1998)	9–10 years	+++	++	Bender. Neuro-paed. exam	S=C S=C	WISC-R Benton	S=C S=C	Reading comprehension	S=C	RSSA; learning problems: socioemotional: hyperactivity: CSR; anxiety: depression: aggression:	S=C ♀: S=C ♂: S<C S=C ♀: S=C ♂: S<C ♀: S=C ♂: S<C ♀: S=C ♂: S<C
Knoester *et al.* (2007a)	5–8 years	++	++	Touwen; MND:	$S=S_1$ $S_1=C$	–	–	–	–	–	–
Knoester *et al.* (2008)	5–8 years	++	++	–	–	RAKIT	$S_1=S$ ♀: $S_1=S$ ♂: $S_1=S$ $S_1<C$ ♀: $S_1<C$ ♂: $S_1<C$	RAKIT; verbal meaning:	$S_1=S$ $S_1<C$	–	–

* Methodological quality: see table II. Int. = Internal validity, Ext. = External validity.

† Neurodevelopmental tests used: Neuro-paed. exam = general neuro-paediatric examination, Bayley = Bayley's Scale of Infant Development (second edition; BSID-II), Stanford-Binet = Stanford-Binet intelligence scale, WPPSI-R = Wechsler Preschool and Primary Scales of Intelligence- Revised

‡ S = study group; infants born following IVF or IVF/ICSI, S1 = study group; infants born following ICSI, C = control group, S = C: no statistically significant differences between study and control group. S > C = Study group performs significantly better than control group, S < C Study group performs significantly worse than control group.

naturally conceived peers. This was especially true for boys. The studies performed by Knoester *et al.* (2007a,2008) could not identify a difference in neuromotor outcome between ICSI and either IVF or naturally conceived children, but cognitive development was slightly worse, i.e. IQ was lower, in singletons born after ICSI than in naturally conceived singletons; this was true for boys and girls.

DISCUSSION

In general, the follow-up studies of good methodological quality showed no consistent differences in neuromotor, cognitive, language and behavioural development between children born following IVF/ICSI and natural conception. We made a great effort to assess methodological quality as a good methodological quality is a prerequisite for generalization of the results of a study. Only 23 papers (39%) met our predefined criteria for good external validity. Thus, our study stresses the need for research with a high methodological quality, i.e. RB studies or truly prospective studies in which all consecutive pregnancies of a hospital or fertility clinic and their naturally conceived controls are followed carefully.

The data from RB studies indicate that IVF/ICSI per se does not increase the risk for CP, but that an increased risk for CP is induced by the association of assisted conception with other risk factors, like preterm birth. Children born following assisted conception are more likely to be born premature and with low birthweight. The increased risk for adverse perinatal outcome cannot solely be attributed to the higher rate of multiplicity following assisted conception. For singletons, the relative risk for preterm and very preterm birth is also found to be increased (Helmerhorst *et al.*, 2004; Jackson *et al.*, 2004). Prematurity and multiple gestations are both strongly associated with the risk of CP (Strömberg *et al.*, 2002; Hvidtjørn *et al.*, 2006). Therefore, these strong risk factors might hide a potential additionally milder effect of IVF/ICSI as causative factor for CP (Dahlquist *et al.*, 2002), if present at all.

It is reassuring that the RB studies did not report an increase in mental retardation or clinically relevant behaviour problems. Since this information is based on large nation-wide registers the chance of a possible non-detected association probably is small. However, a matter of concern is the contradictory findings in the association between IVF/ICSI and epilepsy or convulsions.

To study more subtle differences between IVF/ICSI children and naturally conceived children, such as a minor reduction in IQ, minor neurological or behavioural dysfunction, the controlled studies are most appropriate. In general, these studies did not report an excess of neuromotor, cognitive, language and

behavioural disorders in children born following IVF/ICSI. However, two points should be taken into consideration. First, in the case of neuromotor and cognitive development, most studies used evaluation tools such as the Bayley Scales of Infant Development. These tests have been well standardised and validated as tools for clinical assessment. This means that these instruments are reliable in detecting gross pathology, but they do not evaluate neuromotor and cognitive outcome in a detailed sense. Second, the number of studies of high methodological quality that continued follow-up after infancy is rather limited, while most so-called 'minor' neurodevelopmental disorders are first diagnosed beyond that age.

ICSI

The procedure of ICSI is more invasive in nature than IVF only; natural sperm selection is passed by and spermatozoa with impaired mobility, morphology or genetic abnormalities may be used (Bowen *et al.*, 1998; Sutcliffe *et al.*, 1999, 2001; Knoester *et al.*, 2007a, 2008). In addition, the origin of infertility preceding the fertility therapy is usually different for IVF and ICSI and we do not know how this may affect neurodevelopmental outcome. Overall, the controlled studies included in this review which focused explicitly on follow-up after ICSI showed similar neuromotor, cognitive, language and behavioural outcome after ICSI and natural conception. However, two studies reported a mild IQ reduction in children born after ICSI. The study of Bowen *et al.* (1998) reported significantly lower mental scores in ICSI-boys than IVF- and naturally conceived boys. However, for the generalisability of this study it has to be taken into account that a substantial proportion of infants included in this study were born after the transfer of cryopreserved embryos (Sutcliffe *et al.*, 1998). Recently, Knoester *et al.* (2008) reported lower IQ-scores in 5- to 8-year-old ICSI children when compared to naturally conceived children. This blinded study used a validated test instrument and extensively adjusted for confounders, however, response rate and reasons for non-participation were unclear in the naturally conceived control group. Therefore, ascertainment bias could have emerged: parents who believe that their child is intelligent are probably more willing to let their child cooperate in an intelligence-test, which could have resulted in relatively high IQ-scores in the naturally conceived group. Nevertheless, the results of these two studies warrant more research of high methodological quality and long-term follow-up.

Some studies compared children born following IVF only to children born following IVF with ICSI (Bowen *et al.*, 1998, Leslie *et al.*, 2003; Pinborg *et al.*, 2004; Källén *et al.*, 2005; Hvidtjørn *et al.*, 2006; Knoester *et al.*, 2007a, 2008). The risk of CP between the two treatments did not differ (Pinborg *et al.*, 2004; Källén *et al.*, 2005; Hvidtjørn *et al.*, 2006). The limited data available from controlled studies also

did not suggest that other neurodevelopmental outcome parameters differed for the two treatments (Bowen *et al.*, 1998; Leslie *et al.*, 2003; Knoester *et al.*, 2007a, 2008).

Perspectives for future research

This review demonstrated a clear need for follow-up studies of good methodological quality and with continuation of follow-up after infancy, ideally continuing into adulthood. So far only few studies with solid methodology addressed gender specific vulnerability for neurodevelopmental disorders after IVF or ICSI (Bowen *et al.*, 1998; Levy-Shiff *et al.*, 1998; Knoester *et al.*, 2008). As two of these studies suggest a larger risk for cognitive and emotional problems in boys, we suggest that future studies continue to pay attention to the gender issue.

Another issue which needs evaluation in future research is the long-term effect of cryopreservation of embryos on developmental outcome. Until now very few studies have addressed this problem, while this technique has already been applied on routine base for years. Similar careful and long-term follow-up is warranted for other technologies, such as preimplantation genetic screening or in vitro maturation. We underline the ESHRE Task Force recommendation that with the introduction of new technologies a plan for follow-up should accompany the clinical trial, as the interests of future offspring should be emphasized (ESHRE Task Force on Ethics and Law, 2007).

Conclusion

The majority of studies on neurodevelopmental outcome after IVF/ICSI did not have a robust methodological quality. The exception to this rule was formed by the RB studies. The latter studies suggest that IVF/ICSI per se does not increase the risk for mental retardation or CP. However, the association of assisted conception with risk factors such as multiple gestation and preterm birth does result in an indirect association of IVF/ICSI with CP. The controlled studies which met the criteria of good methodological quality did not show an increase in neuromotor, cognitive, language and behavioural problems in children born after IVF/ICSI. It should be realised that the majority of these studies evaluated outcome in infancy, which precludes a conclusion about the risk of minor neurodevelopmental disorders, which in general first become expressed after infancy.

Part II
The Groningen
ART cohort study

3

Ovarian hyperstimulation and the *in vitro* fertilisation procedure do not influence early neuromotor development, a history of subfertility does.

K.J. Middelburg
M.L. Haadsma
M.J. Heineman
A.F. Bos
M. Hadders-Algra

Fertility Sterility 2010; 93: 544-553.

ABSTRACT

Objective: To evaluate specific effects of ovarian hyperstimulation, the in vitro procedure, and a history of subfertility on neuromotor development at 3 months of age.
Design: Prospective, cohort study.
Setting: University Medical Center Groningen, The Netherlands.
Patient(s): Singletons conceived after controlled ovarian hyperstimulation-IVF/intracytoplasmic sperm injection (COH-IVF; n = 68) or modified natural cycle-IVF/intracytoplasmic sperm injection (MNC-IVF; n = 57), and naturally conceived singletons of subfertile couples (sub-NC; n = 90). Data from a reference population were available (n = 450).
Intervention(s): None.
Main Outcome Measure(s): Quality of general movements (GMs), classified as normal-optimal, normal-suboptimal, mildly abnormal, or definitely abnormal. Definitely abnormal GMs indicate brain dysfunction, mildly abnormal GMs normal but non-optimal brain function.
Result(s): Mildly abnormal and definitely abnormal GMs were observed equally frequently in COH-IVF, MNC-IVF, and sub-NC singletons. The three subfertile groups showed a reduction in GM quality, in particular more mildly abnormal GMs, in comparison with the reference population.
Conclusion(s): Singletons born after IVF (with or without ovarian hyperstimulation) are not at increased risk for abnormal GMs compared with naturally conceived peers of subfertile parents. Mildly abnormal GMs occur more often in infants of subfertile parents than in the general population, suggesting that factors associated with subfertility rather than those related to IVF procedures may be associated with less-optimal early neurodevelopmental outcome. These results need confirmation through replication and follow-up at older ages.

INTRODUCTION

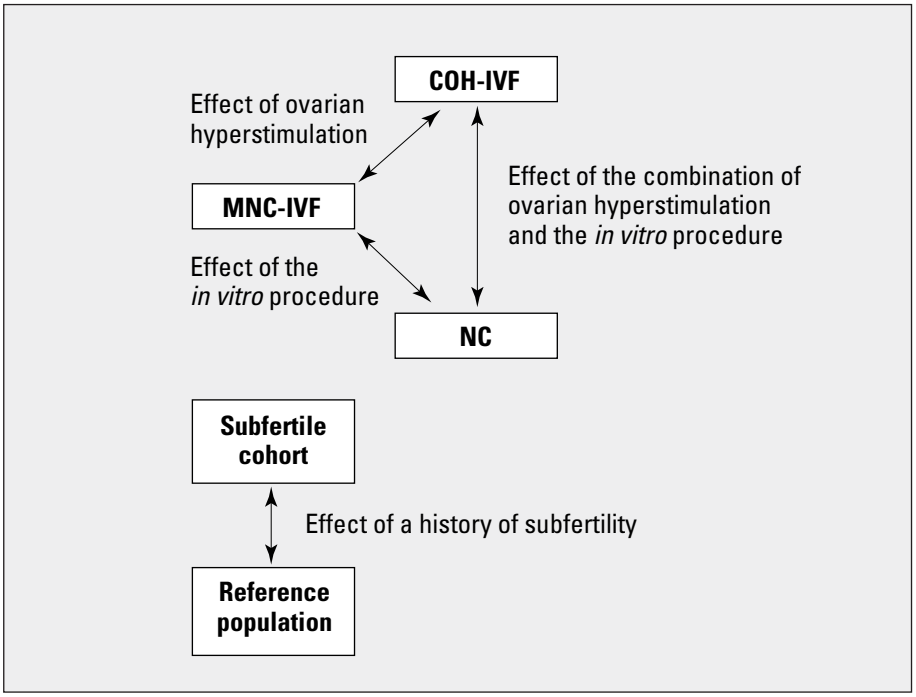
Nowadays a considerable number of children are born after IVF or intracytoplasmic sperm injection (ICSI) (Wright *et al.*, 2008;Andersen *et al.*, 2008), and as a result of these large numbers even subtle changes in the health of these children are of importance to society. It is well known that adverse perinatal outcomes are more common in singletons born after IVF or ICSI than in their naturally conceived peers (Schieve *et al.*, 2002;Helmerhorst *et al.*, 2004;Jackson *et al.*, 2004;Hansen *et al.*, 2008). Therefore, the question arises of whether IVF and ICSI are also associated with adverse neurodevelopmental outcome, either due to the association with low birth weight and prematurity (Bhutta *et al.*, 2002) or as a direct result of fertility treatment or the reason for subfertility. Recently a systematic review on neurodevelopmental outcome of children born after IVF/ICSI concluded the following. First, results from large, register-based studies suggest that IVF/ICSI per se does not increase the risk of severe neurodevelopmental handicaps, like cerebral palsy (Middelburg *et al.*, 2008). However, indirectly (i.e., by means of the association with preterm birth and plurality), IVF/ICSI is related to an increased risk of cerebral palsy (Strömberg *et al.*, 2002;Klemetti *et al.*, 2006;Hvidtjørn *et al.*, 2006;Källén *et al.*, 2005). Second, controlled studies that met criteria for good methodologic quality, such as blinded assessors, prospective design, high follow-up rates, adequate correction for confounders, and the use of validated neurodevelopmental test instruments, did not show an increase in neurodevelopmental problems in IVF/ICSI children (Middelburg *et al.*, 2008). However, here it should be realised that the studies in infancy did provide reassuring results but used relatively gross measures to document outcome (Middelburg *et al.*, 2008). Valid evidence in school-aged children is still scarce (Knoester *et al.*, 2007b;Knoester *et al.*, 2008;Levy-Shiff *et al.*, 1998).

It is conceivable that one or more components of the IVF procedure induce change in embryo development. Hypothetical points of concern are, for instance, ovarian hyperstimulation, the IVF/ICSI procedure itself, and consequences of vanishing twins (Olivennes *et al.*, 1993;Jackson *et al.*, 2004;Kapiteijn *et al.*, 2006;Pinborg *et al.*, 2005;Pinborg *et al.*, 2007;Griesinger *et al.*, 2008). In addition, background factors associated with the IVF/ICSI procedure, such as a history of subfertility or increased parental age, may contribute (Draper *et al.*, 1999;Lambert, 2003;Thomson *et al.*, 2005;Sutcliffe and Ludwig, 2007).

In this prospective, cohort study, we addressed the question of whether neurodevelopmental outcome is related to ovarian hyperstimulation, the in vitro procedure itself, a combination of these two factors, or a history of subfertility. To this end, we made four comparisons (figure 1). First, to study the effect of ovarian hyperstimulation, we compared the quality of general movements (GMs)

in infants born after "conventional" controlled ovarian hyperstimulation IVF (COH-IVF) with that in infants born after IVF in the modified natural cycle (MNC-IVF). In MNC-IVF the aim is to use the one follicle that naturally develops to dominance. Therefore, medication use is minimal and starts only after follicular dominance has developed (Rongières-Bertrand *et al.*, 1999). Second, to study the effect of the in vitro procedure, we compared the quality of GMs in infants born after MNC-IVF with that in naturally conceived (sub-NC) infants born to subfertile couples who tried to conceive for at least 1 year and waited for their fertility workup or treatment. Third, by comparing COH-IVF and sub-NC infants, we studied the combined effect of ovarian hyperstimulation and the in vitro procedure. This information is most valuable for subfertile couples who consider IVF treatment, because in general IVF is applied after COH. Finally, the three cohorts (COH-IVF, MNC-IVF, and sub-NC) were taken together to form one subfertile cohort. The quality of GMs in this subfertile cohort was then compared with that in a reference population, to study the effect of subfertility itself and subfertility-related parental characteristics (figure 1). The differentiation of the effects of ovarian hyperstimulation, the in vitro procedure, and subfertility on neurodevelopmental outcome is a unique aspect of our study.

Figure 1 - The effects of ovarian hyperstimulation, the in vitro procedure itself, and a history of subfertility are studied by four different comparisons. The subfertile cohort is formed by the COH-IVF, MNC-IVF and sub-NC groups taken together.



The detection of subtle differences in neurodevelopmental outcome requires sensitive tests. At early age, the assessment of the quality of GMs is such a sensitive method. For instance, it allowed for the detection of subtle differences in neurodevelopmental outcome of healthy term infants who received formula with or without long-chain polyunsaturated fatty acids (Bouwstra *et al.*, 2003b). General movements are spontaneous, nonvoluntary movements of the fetus and young infant involving all parts of the body. They persist until voluntary movements gradually emerge, which is from approximately 4 months after term onward. The quality of GMs reflects the condition of the central nervous system at a young age (Prechtl, 1990). A functionally intact nervous system produces movements characterized by complexity, variation, and—to a lesser extent—fluency (table I;Hadders-Algra *et al.*, 2004). The advantage of assessing neurodevelopmental outcome at early age is the relatively small impact of postnatal factors, in particular social conditions, on developmental outcome, which allows for a closer linkage of findings to early ontogenetic events. The disadvantage of using outcome at early age is that neurobehavioural condition at early age is related to a limited extent to outcome at school age (Hadders-Algra, 2002).

METHODOLOGY

Recruitment

From March 2005 to December 2006 infants of subfertile couples were recruited at the Department of Reproductive Medicine of the University Medical Center Groningen, a tertiary care center in the Netherlands. All couples with a singleton pregnancy after IVF/ICSI were invited for participation in a longitudinal study on neurodevelopmental outcome of IVF/ICSI children. This resulted in two groups, the first consisting of singletons born after "conventional" IVF/ICSI in the COH cycle (COH-IVF) and the second consisting of singletons born after IVF/ICSI in the modified natural cycle (MNC-IVF). Details on treatment protocol and procedures in MNC-IVF have previously been described by Pelinck *et al.* (Pelinck *et al.*, 2007;Pelinck *et al.*, 2008). Excluded from the study were infants born after treatment with cryopreserved or donated oocytes or embryos.

For participation in the sub-NC control cohort all couples were invited who achieved a singleton pregnancy while on the waiting list for fertility evaluation or treatment during the study period. These couples had been subfertile for at least 1 year. This cohort was chosen as a control group because we expected that parental characteristics like parity, age, and possibly other unknown factors would resemble the characteristics of IVF couples more than those of couples in the general

population. Excluded from the sub-NC group were couples with pregnancies resulting from any other form of assisted reproduction (e.g., ovulation induction and/or insemination). Parents of infants in the COH-IVF, MNC-IVF, and sub-NC groups were invited to participate during the third trimester of pregnancy.

Information on a reference population was available; it was recruited at six child welfare clinics in the northern part of the Netherlands (Bouwstra *et al.*, 2009). In 2001 all parents of 12-week-old infants visiting the child welfare clinic for routine general health care (n = 605) were invited to participate. Parents of 70 infants (12%) refused to participate, and videos of another 80 (13%) infants could not be adequately assessed owing to insufficient quality of the recordings. An adequate recording of GMs was made in 455 infants (75%). The reference population was representative of the Dutch general population for birth weight, rate of preterm birth, and maternal age at birth (CBS; StatLine databank). For the present study 5 multiples were excluded, so that information on 450 singletons was available. Information on the use of any form of assisted reproduction in the reference population was not available.

The ethics committee of the University Medical Center Groningen approved the study design, and all parents provided written, informed consent for participation of their infants in the study.

Demographics

Information on the prenatal, perinatal, and neonatal periods, parental characteristics, and socioeconomic conditions were collected on standardised charts at the first follow-up assessment. For the COH-IVF, MNC-IVF, and sub-NC groups this was at the assessment that was scheduled 2 weeks after term; for the reference population this was at the assessment at 3 months after term. Extra information from midwifes and gynecologists was obtained when complications had occurred during pregnancy or birth or when information was incomplete. Details on treatment procedures, fertility diagnosis, and time to pregnancy were obtained from medical records in the COH-IVF, MNC-IVF, and sub-NC groups.

General Movements

Quality of GMs is classified into four different categories: two normal variants (normal-optimal [NO] and normal-suboptimal [SO]) and two abnormal variants (mildly abnormal [MA] and definitely abnormal [DA]); criteria are shown in Table I (Hadders-Algra *et al.*, 2004). The reliability and validity of the GM method is good (Hadders-Algra *et al.*, 2004;Heineman and Hadders-Algra, 2008). Multiple studies of high-risk infants have demonstrated that DA GMs at 2–4 months after term accurately predict cerebral palsy and that MA GMs are associated with minor

neurologic dysfunction and behavioural problems at school age (Hadders-Algra *et al.*, 2004;Hadders-Algra and Groothuis, 1999;Groen *et al.*, 2005). However, it should be kept in mind that a mildly abnormal quality of GMs as an isolated risk factor is a poor predictor of an infant's neurodevelopmental outcome. In other words, MA GMs do not imply a mildly abnormal nervous system but rather a normal but non-optimally functioning nervous system (table I). Suboptimal GM quality can be regarded as the typical GM quality shown by most infants born at term (Bouwstra *et al.*, 2003b). Normal-optimal GMs are relatively rare; their occurrence is associated with breastfeeding (Bouwstra *et al.*, 2003a).

Table I - Classification of the quality of General Movements, Table adapted from Hadders-Algra et al., 2004.

Classification of GM-quality		Complexity[a]	Variation[b]	Fluency[c]	Corresponding brain function
Normal-optimal GMs	(NO)	+++	+++	+	excellent
Normal-suboptimal GMs	(SO)	++	++	-	typical
Mildly abnormal GMs	(MA)	+	+	-	non-optimal
Definitely abnormal GMs	(DA)	-	-	-	dysfunction

Note: Originally published in Clinical Rehabilitation.
[a] *GM-complexity* = spatial variation;
[b] *GM-variation* = temporal variation
[c] *GM-fluency* = fluency of the movements
+++ = abundantly present, ++ = sufficiently present, + = present to a limited extend, - = virtually absent.

General movement quality in COH-IVF, MNC-IVF, and sub-NC singletons was assessed at 2 weeks and 3 months after term. The reference population had only been assessed at age 3 months. Spontaneous movements in the supine position were videotaped for 5–10 minutes. The aim was to record the infant's motility in an awake, active, and not-crying behavioural state because these conditions influence movement quality.

Two specialized assessors (K.J.M. and M.H.A.) independently analysed all video recordings of infants born after COH-IVF, MNC-IVF, and sub-NC. Different scores were discussed until consensus was met. Interobserver reliability of the original scores was good (determined on a random sample of 70 videos: $\kappa = 0.82$, 91% agreement; similar agreement for 2 weeks and 3 months). All videos in the reference population also had been analysed by two investigators, one being the same as for the cohort study (M.H.A.; interobserver reliability in the

reference population: $\kappa = 0.82$, 90% agreement). Assessors were blind for mode of conception of infants in the cohort study (COH-IVF, MNC-IVF, or sub-NC). The reference population was recruited previously and therefore assessors were not blind to group status ("child welfare clinics visitor") of these infants.

Statistical Analyses

This study is part of a longitudinal study on neurodevelopmental outcome after IVF/ICSI. Power calculation of the longitudinal study is based on prevalence of minor neurologic dysfunction; however, this can only be assessed from 4 months onward. We performed a post hoc power calculation for GM quality. The prevalence of abnormal GMs in the general population was estimated at 25% (Bouwstra *et al.*, 2003b). In this case, a sample of 58 infants leads to 80% power to detect a doubling of the prevalence of abnormal GMs.

Chi-square and Fisher exact tests were used to assess distribution of abnormal (including MA and DA) GMs and DA GMs. The influence of ovarian hyperstimulation and the in vitro procedure on abnormal or DA GMs at the age of 3 months after term was analysed using multiple logistic regression. Dummy variables were created for mode of conception. The reference population was included in a second run of logistic regression analysis to study the effect of a history of subfertility. Variables for which groups differed at $P < .05$ were entered into the multivariate analysis to correct for their influence on GM quality. In addition, gestational age was included in the multivariate analyses, because we know from the literature that it is an important predictor for quality of GMs. Adjusted odds ratios (ORs) were then calculated for DA GMs in the different groups. In a separate analysis we explored whether children born after ICSI (compared with IVF) were more likely to show abnormal (MA + DA) or DA GMs, given that this procedure is more invasive in nature and impaired spermatozoa may be used, which may have consequences for neurodevelopmental outcome. Finally, we explored whether interaction between gender and conception mode was a predictor for abnormal (MA + DA) or DA GMs because vulnerability for impaired neurodevelopmental outcome may be gender specific; some studies have reported differences in neurodevelopmental outcomes between boys and girls (Bowen *et al.*, 1998;Levy-Shiff *et al.*, 1998;Knoester *et al.*, 2008;Te Velde *et al.*, 1998;Knoester *et al.*, 2007a). Probability values $\leq .05$ were considered significant. Correction for multiple testing was not performed in the analysis of patient characteristics, so as to be transparent concerning differences in characteristics between the groups. Statistical analyses were performed using SPSS 14.0 for Windows (SPSS, Chicago, IL).

Table II - Characteristics of parents and infants born following controlled ovarian hyperstimulation IVF (COH-IVF), modified natural cycle IVF (MNC-IVF) and naturally conceived controls born to subfertile parents (NC).

Characteristics	COH-IVF	MNC-IVF	NC
	n = 68	*n = 57*	*n = 90*
Male gender; n (%)	36 (53%)	27 (47%)	46 (51%)
First born, n (%)	47 (69%)	38 (67%)	55 (61%)
Gestational characteristics:			
Vanishing twins, n (%)	8 (12%)*/**	1 (2%)*	0 (0%)**
Pregnancy induced hypertension, n (%)	8 (12%)	3 (5%)	15 (17%)
Signs of fetal distress[a], n (%)	20 (29%)	16 (28%)*	40 (44%)*
Forceps/vacuum extraction, n (%)	6 (9%)	7 (12%)	11 (12%)
Caesarean section, n (%)	17 (25%)	8 (14%)	24 (27%)
Birth characteristics:			
Gestational age in weeks; median (range)	39.4 (33-42)*	40.1 (35-43)	40.0 (30-43)*
Preterm birth (< 37 weeks); n (%)	7 (10%)	6 (11%)	7 (8%)
Birth weight in grams; median (range)	3378 (1980-4700)*	3400 (2170-4680)	3565 (1150-4710)*
Low birth weight (< 2500 gram); n (%)	3 (4%)	4 (7%)	5 (6%)
Small for gestational age[b]; n (%)	0 (0%)	3 (5%)	2 (2%)
Missing values for various neonatal, parental and fertility variables	< 4	< 3	< 5
Neonatal characteristics:			
Apgar score 5 min < 7, n (%)	0 (0%)	0 (0%)	1 (1%)
NICU admission, n (%)	1 (2%)	2 (4%)	7 (8%)
Breastfed for > 6 weeks, n (%)	30 (46%)	26 (46%)	42 (48%)
Parental characteristics:			
Maternal age at conception; median (range)	32.5 (26-41)	32.8 (25-37)	33.2 (22-40)
Paternal age at conception; median (range)	35.7 (28-56)	34.4 (28-48)	35.4 (25-53)
Smoking during pregnancy, n (%)	7 (10%)	7 (12%)	9 (10%)
Parental socioeconomic status:			
Education level mother (high[c]), n (%)	22 (32%)	22 (39%)	41 (46%)
Education level father (high[c]), n (%)	29 (45%)	19 (34%)	33 (37%)
Fertility factors:			
Intracytoplasmic sperm injection; n (%)	43 (63%)	29 (51%)	-
Time to pregnancy in years; median (range)	4.1 (0-13)***	3.8 (0-13)**	2.1 (0-11)***/**
Corrected age at examination (in weeks):			
Two weeks; median (range)	2.4 (0-5)	2.7 (1-6)	2.6 (1-5)
Three months; median (range)	13.0 (11-15)	13.1 (11-15)	13.1 (12-17)

*Note: Mann-Whitney U test and Chi-square test were used; * P < .05, ** P < .01, *** P < .001.*

[a] Signs of fetal distress denoted by meconium stained amniotic fluid and/or cardiotocographic signs and/or acidosis.

[b] Birth weight for gestational age is < -2 standard deviation scores compared to a Dutch reference population (Dutch reference tables, Perinatal Registration Netherlands).

[c] University education or vocational colleges.

RESULTS

Infant and Parental Characteristics

In the inclusion period 89 infants were born after COH-IVF, 79 after MNC-IVF, and 143 after a natural conception; of these infants, 68 (76%), 57 (72%), and 90 (63%), respectively, were included in the study. Characteristics of participants and nonparticipants were compared, and it turned out that participation was nonselective. Gender, birth weight, gestational age, firstborn infants, the percentage of infants born preterm or small for gestational age (–2 SD scores), admission to the neonatal intensive care unit, parental educational level, and time to pregnancy were similar for participants and nonparticipants (data not shown).

Table III - Characteristics of the subfertile cohort and the reference population.

	Subfertile cohort[a]	**Reference population**
	n = 215	***n = 450***
Male gender; n (%)	109 (51%)	212 (48%)[d]
First born, n (%)**	140 (65%)	197 (44%)[c]
Birth weight in grams, mean (range)	3453 (1150-4710)	3459 (850-5432)[d]
Low birth weight (< 2500 gram), n (%)	12 (5.6%)	29 (6.6%)[d]
Gestational age (in weeks), median (range)	40.0 (30-43)	40.0 (29-42)[c]
Preterm birth (< 37 weeks), n (%)	20 (9.3%)	31 (6.9%)
Forceps/ vacuum extraction, n (%)	24 (11%)	53 (12%)
Caesarean section, n (%)**	49 (23%)	54 (12%)[d]
Breastfed for > 6 weeks, n (%)	98 (47%)[c]	234 (53%)[d]
Parental characteristics:		
Maternal age at conception, mean (range)**	32.7 (22-41)	30.5 (15-46)[e]
Paternal age at conception, mean (range)**	36.0 (25-56)[c]	32.8 (16-63)[e]
Smoking during pregnancy (mother), n (%)	23 (11%)	61 (14%)
Parental socioeconomic status:		
Education level mother (high[b]), n (%)	85 (40%)	147 (33%)[c]
Education level father (high[b]), n (%)	81 (38%)[c]	162 (36%)[c]
Corrected age at examination (in weeks):		
Three months, median (range)**	13 (11-17)	14 (7-16)[e]

*Note: Students t-test and χ^2 test were used; * P < .05, ** P < .001.*

[a] The subfertile cohort consists of the infants born following COH-IVF, MNC-IVF and NC taken together.

[b] University education or vocational colleges.

[c] ≤ 5 missing values for this variable.

[d] ≤ 10 missing values for this variable.

[e] ≤ 52 missing values for this variable.

An exception was maternal age at conception: nonparticipating sub-NC mothers were significantly younger than participating sub-NC mothers (P =.03).
Characteristics of the singletons born after COH-IVF, MNC-IVF, and sub-NC are presented in table II. For the majority of parental, gestational, birth, and neonatal variables the groups were comparable. However, sub-NC pregnancies were more frequently complicated by signs of fetal distress than IVF pregnancies—a difference that reached statistical difference for the comparison with MNC-IVF (P =.05). Birth weight was significantly higher after natural conception than after COH-IVF (P =.02), whereas birth weight of MNC singletons fell in between of that of the other two groups. Similarly, gestational age in the sub-NC group was significantly higher than in the COH-IVF group (P =.02). Time to pregnancy in the sub-NC group was significantly shorter than in the IVF groups (COH-IVF, P <.001; MNC-IVF, P =.002).

The gestational, neonatal, and parental characteristics of the three groups (COH-IVF, MNC-IVF, and sub-NC) were largely comparable; we pooled these three groups to form a subfertile cohort for the comparison with the reference population. On average, subfertile parents were older (P <.001), and their infants were more often first born (P <.001) or born by cesarean section (P <.001). Infants in the reference population were slightly older at testing than infants in the cohort study (P <.001) (table III).

Quality of GMs

At 2 weeks after term, 67 infants (99%) born after COH-IVF, 57 (100%) after MNC-IVF, and 89 sub-NC infants (99%) were videotaped. Four recordings (1 MNC-IVF and 3 sub-NC infants) could not be assessed because the infants were continuously crying. The distribution of the quality of GMs at the age of 2 weeks is presented in table IV. The frequency of NO, SO, and abnormal GMs (MA + DA) was similar for infants born after COH-IVF, MNC-IVF and sub-NC infants. Infants born after COH-IVF tended to show more often DA GMs (6%) than MNC-IVF (0) and sub-NC infants (1%) at 2 weeks, but these differences did not reach statistical significance (P =.12, P =.17).

At 3 months after term, 68 infants (100%) born after COH-IVF, 56 (98%) after MNC-IVF, and 88 sub-NC infants (98%) were videotaped. Table IV shows the distribution of the quality of GMs at 3 months. The rate of NO, SO, abnormal (MA + DA), and DA movements did not differ significantly between the COH-IVF, MNC-IVF, and sub-NC groups.

A remarkable finding was that GM quality in the subfertile cohort (COH-IVF, MNC-IVF, and sub-NC taken together) was reduced compared with that in the reference population (table IV). In the subfertile cohort 3% of infants showed NO GMs, compared with 14% in the reference population, whereas 42% and 28% of

Table IV - Distribution of GM-quality.

	GM-quality[a]			
	normal		abnormal	
	NO	SO	MA	DA
GM-quality at the age of two weeks				
COH-IVF [b]	2 (3%)	41 (61%)	20 (30%)	4 (6%)
MNC-IVF [b]	3 (5%)	32 (57%)	21 (38%)	0 (0%)
NC [b]	5 (6%)	47 (55%)	33 (38%)	1 (1%)
GM-quality at the age of three months				
COH-IVF [b]	2 (3%)	36 (53%)	30 (44%)	0 (0%)
MNC-IVF [b]	3 (5%)	35 (63%)	18 (32%)	0 (0%)
NC [b]	2 (2%)	46 (52%)	39 (44%)	1 (1%)
GM-quality at the age of three months in comparison to a reference population				
Subfertile cohort [c] *	7 (3%)	117 (55%)	87 (41%)	1 (0.5%)
Reference population*	62 (14%)	263 (58%)	109 (24%)	16 (4%)

*Note: χ^2 test for trend was used to compare quality of General Movements between the conception groups; * $P < .001$.*

[a] NO = normal-optimal, SO = normal-suboptimal, MA = mildly abnormal, DA = definitely abnormal.

[b] COH-IVF = controlled ovarian hyperstimulation IVF, MNC-IVF = modified natural cycle IVF, NC = naturally conceived infants of subfertile parents.

[c] The subfertile cohort is formed by the COH-IVF, MNC-IVF, and NC group taken together.

infants, respectively, showed abnormal (MA + DA) GMs. Conversely, DA GMs were observed more frequently in the reference population: 4% compared with 0 in the subfertile cohort. Overall, GM quality was statistically significantly reduced in the subfertile cohort compared with the reference population (χ^2 test for trend, P <.001).

Results of logistic regression analysis for GM quality at the age of 3 months are shown in table V. The different conception methods were used as explanatory factors. The adjusted OR for abnormal GMs in the COH-IVF and MNC-IVF groups compared with the sub-NC group was, respectively, 0.91 (95% confidence interval [CI], 0.46–1.81) and 0.61 (95% CI, 0.29–1.28). When MNC-IVF was used as the indicator the adjusted OR for abnormal GMs in the COH-IVF group was 1.49 (95% CI, 0.70–3.18). This means that an abnormal quality of GMs could not be explained by the in vitro procedure or by ovarian hyperstimulation or by the combination of ovarian hyperstimulation and the in vitro procedure.

In the secondary logistic regression analysis, the reference population was used as the indicator. This analysis showed that being a member of the subfertile cohort was indeed associated with a higher prevalence of abnormal general movements (adjusted OR 1.83 [95% CI, 1.25–2.68]; table V.). Missing values for cesarean section or age of examination in the reference group for children with DA

GMs resulted in a reduction of the number of available cases for the multivariate analysis. For this reason, we tested models with and without these covariates included; the analyses revealed that in both models the contribution of group status to GM quality was similar (we reported the model with the least covariates). The lower occurrence of DA GMs in the subfertile cohort was no longer statistically significant after correction for confounders in the logistic regression analysis (adjusted OR 0.14 [95% CI, 0.02–1.12]).

Further analysis of covariates could not identify ICSI or interaction between gender and conception method as significant predictors for abnormal (MA + DA) or DA GMs.

Table V - Logistic regression analyses of contribution of IVF-method and a history of subfertility to the rate of abnormal (MA + DA) and definitely abnormal (DA) GMs at the age of three months.

Abnormal (MA[a] + DA[a]) GMs at the age of three months

Covariate	Indicator	OR [95% CI][d]	Adjusted OR [95% CI][d]
COH-IVF[b]	NC[b]	0.95 [0.50-1.79]	0.91 [0.46-1.81][e]
MNC-IVF[b]	NC[b]	0.57 [0.28-1.15]	0.61 [0.29-1.28][f]
COH-IVF[b]	MNC-IVF[b]	1.67 [0.80-3.48]	1.49 [0.70-3.18][g]
Subfertile cohort[c]	Reference population	1.84 [1.31-2.60]	1.83 [1.25-2.68][h]

Definitely abnormal (DA) GMs at the age of three months[j]

Covariate	Indicator	OR [95% CI][d]	Adjusted OR [95% CI][d]
Subfertile cohort[c]	Reference population	0.13 [0.02-0.98]	0.14 [0.02-1.12][i]

[a] MA = mildly abnormal, DA = definitely abnormal.
[b] COH-IVF = controlled ovarian hyperstimulation IVF, MNC-IVF = modified natural cycle IVF, NC = naturally conceived infants of subfertile parents
[c] The subfertile cohort is formed by the COH-IVF, MNC-IVF, and NC group taken together.
[d] Odds ratio/adjusted odds ratio with 95% confidence interval.
[e] Corrected for gestational age, birth weight, vanishing twins, and time to pregnancy.
[f] Corrected for gestational age, signs of fetal distress, and time to pregnancy.
[g] Corrected for gestational age, and vanishing twins.
[h] Corrected for gestational age, primiparity, caesarean section, maternal age, and age at examination.
[i] Corrected for gestational age, primiparity, maternal age.
[j] Logistic regression analysis for factors predicting DA GMs in the COH-IVF, MNC-IVF and NC group separately was not possible since only 1 child showed DA movements.

DISCUSSION

The present study demonstrated no association between mode of conception and quality of GMs. At the ages of 2 weeks and 3 months, movement quality was similar for singletons born after COH-IVF, MNC-IVF and their naturally conceived peers born to subfertile parents. However, being born to subfertile parents was associated with a reduced GM quality at the age of 3 months in comparison with a reference population. This indicates that neither ovarian hyperstimulation nor the in vitro procedure but rather factors associated with subfertility affect early neurodevelopmental outcome.

Strengths and Limitations

One of the strengths of this study is the subfertile control group. We selected this control group so as not to overestimate a potential effect of fertility treatment. In this manner, we composed a control cohort that closely resembled the study group, for instance regarding maternal age and parity. Nevertheless, the sub-NC group was not perfectly similar to the IVF groups: time to pregnancy was significantly longer in the IVF groups, and other factors associated with the ability to conceive naturally may have been different. Surprisingly, we observed a high percentage of infants with signs of fetal distress in our subfertile sub-NC group. Likewise, the sub-NC pregnancies were frequently complicated by pregnancy-induced hypertension, and sub-NC children were relatively often born by cesarean section or admitted to neonatal intensive care. This may have been a chance finding but may also be the result of an increased risk in pregnancies of subfertile couples, and obstetric care might have been less intensive in sub-NC than in IVF pregnancies. General movement quality in sub-NC children may be reduced as a consequence of the high percentage of fetal distress in this group, which could conceal a reduction in GM quality in IVF infants.

The prospective design of this study, in which couples were included during pregnancy, and the policy of the Department of Reproductive Medicine to collect pregnancy and birth details of all patients allowed us to evaluate selection bias. This turned out to be virtually absent. With the initial prenatal enrollment of 63%–76% of eligible infants and with low postnatal attrition (0–2%), we assume to have a representative sample.

We used a reliable and standardised test instrument that allowed us to study subtle differences in neurodevelopmental outcome early in life. The clinical value of GM assessment for early detection of developmental disorders in high-risk populations is well established (Hadders-Algra *et al.*, 2004;Hadders-Algra, 2004). In these populations, DA GMs at 3 months indicate a high risk for the development

of cerebral palsy. Mildly abnormal GMs are associated with development of minor neurological dysfunction, attention deficit–hyperactivity disorder, and aggressive behaviour at early school age (Hadders-Algra and Groothuis, 1999) and coordination problems, fine manipulative disability, and psychiatric morbidity at the age of 9–12 years (Groen *et al.*, 2005;Hadders-Algra *et al.*, 2008). General movement assessment has also been recognised as a research tool to assess the effect of prenatal, perinatal, and early postnatal conditions on the young nervous system (Bouwstra *et al.*, 2003b). However, it should be realised that no data are available on the predictive value of GM quality in the general population. This means that we do not know what the significance of the current findings is for the children's long-term outcome. But because fertility problems are steadily rising in society, an increase in the occurrence of, for instance, minor neurological dysfunction, attention deficit–hyperactivity disorder, and coordination problems would be relevant for the general population. Our results certainly warrant further follow-up because they may imply that children of subfertile parents are at increased risk for various neurodevelopmental disorders.

A limitation in the design of our study is that the medication used in MNC-IVF, although minimal, could have caused an overestimation of the effect of IVF or an underestimation of the effect of COH. In the interpretation of the results of our study we were, however, not hampered by this minor confounding of MNC, because assisted reproduction was not associated with reduced GM quality.

The secondary analysis (the comparison between the total subfertile cohort and the reference population) was limited by the fact that the assessors were aware of the group status of the reference population because of recruitment of this group at child welfare clinics. Furthermore, our reference population was different in terms of, for example, parity, maternal age, and percentage of cesarean sections. For this reason, we corrected for these differences in the multivariate analyses. No information was available on mode of conception of the reference children, but it is rather likely that some of the children were born after assisted reproduction. This may imply that actual differences in GM quality between children of fertile and subfertile couples are larger than indicated in the present study. Last, we should emphasize that this study has limited power to detect developmental disorders of low incidence (i.e., the occurrence of DA GMs or small differences between the groups), owing to the relatively small sample sizes in the subgroups.

Infants of Subfertile Couples

Previous studies have reported an increased risk of obstetric complications and adverse perinatal outcome in subfertile couples (Draper *et al.*, 1999;Pandian *et al.*, 2001;Thomson *et al.*, 2005). The literature is, however, inconclusive as to the

etiology of this increased risk. Some studies suggest that fertility treatment itself or parameters for ovarian stimulation do not influence perinatal outcome, but rather the subfertility per se Thomson *et al.*, 2005;Romundstad *et al.*, 2008;Griesinger *et al.*, 2008). Other studies, however, suggest that fertility treatment or hormonal stimulation per se influence perinatal outcome (Kallen *et al.*, 2005;Wennerholm *et al.*, 2000;Wang *et al.*, 2005;Kapiteijn *et al.*, 2006;De Geyter *et al.*, 2006). Our study did not distinguish between an effect of COH or the in vitro procedure itself on perinatal outcome, which may be related to the relatively small sample size.

The reduced GM quality in infants of subfertile couples found in our study can not be attributed solely to an increase in adverse perinatal outcome, because gestational age and being born by cesarean section were both included as covariates into the logistic regression analysis. In addition, differences in breastfeeding practice could not explain the differences in GM quality between the subfertile cohort and the reference population because the rate of infants who were breastfed for at least 6 weeks was similar in both groups. Other pathophysiologic reasons for the increased rate of abnormal GMs in the subfertile cohort might be genetic makeup or altered hormonal conditions. It is conceivable that a non-optimal genetic or hormonal condition results not only in subfertility but also in a less optimal neurodevelopmental outcome.

To date, three other studies have evaluated neurodevelopmental outcome in IVF children in comparison with children of subfertile couples. Sun *et al.* (Sun *et al.*, 2007) addressed the prevalence of epilepsy and febrile seizures and made a distinction between subfertile couples who conceived naturally and those who conceived by fertility treatment. They reported that children of treated subfertile couples showed an increased risk of epilepsy compared with children of fertile couples, whereas the risk in children of untreated subfertile couples was elevated but not to a statistically significant level. This result may seem at variance with our findings. It should be realised, however, that epilepsy and febrile seizures are specific neurologic entities, whereas the quality of GMs is a parameter of general neurologic function. Wagenaar *et al.* reported on school functioning, behaviour, and socioemotional functioning of 9–18-year old IVF children (Wagenaar *et al.*, 2008b;Wagenaar *et al.*, 2008a). School functioning was similar; however, IVF children were found to show fewer problem behaviour and attention problems but more withdrawn or depressed behaviour compared with children of subfertile parents. The investigators speculate that these findings might be caused by higher hypothalamic–pituitary–adrenal axis activity because of changes in endocrine and metabolic processes by IVF (Wagenaar *et al.*, 2008b).

It would be of great interest to elucidate the mechanisms behind the association of subfertility and perinatal or developmental outcome in future

studies. This could lead to identification of possible at-risk subfertile couples and customised obstetric or child welfare care for these subgroups of patients.

In summary, in this prospective cohort study we found no differences in GM quality between infants born after COH-IVF, MNC-IVF and infants of couples who conceived naturally during fertility evaluation or while on the waiting list for fertility treatment, indicating no effect of ovarian hyperstimulation or the in vitro procedure on neurodevelopmental outcome early in infancy. However, we found a substantial reduction in GM quality in a subfertile cohort compared with a reference population at 3 months after term, suggesting that factors associated with subfertility might negatively affect neurodevelopmental outcome. Both results should be interpreted with caution because long-term follow-up and confirmation of these results in studies with a larger sample size is necessary to draw firm conclusions.

Part II
The Groningen
ART cohort study

4

Ovarian hyperstimulation and the *in vitro* procedure do not affect neurological outcome in infancy

K.J. Middelburg
M.J. Heineman
A.F. Bos
M. Pereboom
V. Fidler
M. Hadders-Algra

Human Reproduction 2009; 24: 3119-3126.

ABSTRACT

Background: Due to the growing number of children born following assisted reproduction technology, even subtle changes in the children's health and development are of importance to society at large. The aim of the present study was to evaluate the specific effects of ovarian hyperstimulation and the *in vitro* procedure on neurological outcome in 4–18-month-old children.
Methods: In this prospective assessor-blinded cohort study, we included singletons born following controlled ovarian hyperstimulation in vitro fertilisation (IVF) or intracytoplasmic sperm injection (ICSI) (COH-IVF; n = 68) or modified natural cycle-IVF/ICSI (MNC-IVF; n = 57) or naturally conceived singletons of subfertile couples (sub-NC; n = 90). Children were assessed with standardised, age-specific and sensitive neurological assessments (TINE and Hempel assessment) at 4, 10 and 18 months. Neurological examination resulted in a neurological optimality score (NOS), a fluency score and a clinical neurological classification. Fluency of movements is easily affected by neurological dysfunction and is therefore a sensitive measure for minimal changes in neuromotor development.
Results: The NOS and the fluency score were similar in COH-IVF, MNC-IVF and sub-NC children. None of the children showed major neurological dysfunction and rates of minor neurological dysfunction at the three ages were not different between the three conception groups.
Conclusions: We found no effects of ovarian hyperstimulation or the *in vitro* procedure itself on neurological outcome in children aged 4–18 months. The findings of our study are reassuring, nevertheless it should be kept in mind that subtle neurodevelopmental disorders may emerge when children grow older. Continuation of follow-up in older and larger groups of children is therefore still needed.

INTRODUCTION

The number of children born following assisted reproductive technology (ART) will become substantial in the coming decades. Worldwide, registers have reported increases in the percentage of children born following ART, e.g. in Scandinavia, already up to 4% of children are born following ART (Andersen *et al.*, 2008; Wright *et al.*, 2008).

Due to the growing number of ART-conceived children, even minimal changes in the children's health and development are of importance to society at large. Up to now, results of most developmental studies have been reassuring (reviewed by Sutcliffe and Ludwig, 2007; Middelburg *et al.*, 2008). Nevertheless, many studies are hampered by methodological shortcomings, such as non- or partially blinded observers, differences in the recruitment of study and control children, high attrition rates and the use of neurodevelopmental tests not sensitive enough to detect subtle differences (Middelburg *et al.*, 2008). Furthermore, evidence suggests that singletons born following ART are at increased risk for preterm birth and low birthweight (Helmerhorst *et al.*, 2004; Jackson *et al.*, 2004). As the latter conditions are related to impaired development (Bhutta *et al.*, 2002; Moster *et al.*, 2008), this finding has generated great concern.

In theory, various components of the ART procedure may change embryo development and in that way influence health or development of the conceived child. Suggested points of concern are the effects of laboratory procedures involved with in vitro fertilisation (IVF) or intracytoplasmic sperm injection (ICSI), the effects of ovarian hyperstimulation (bypassing natural selection of the dominant follicle and possibly causing diminished endometrial receptivity by supraphysiological estradiol levels) and consequences of vanishing twins (Olivennes *et al.*, 1993; Draper *et al.*, 1999; Jackson *et al.*, 2004; Pinborg *et al.*, 2005; Kapiteijn *et al.*, 2006; Sutcliffe and Ludwig, 2007; Griesinger *et al.*, 2008). But, parental characteristics associated with subfertility may also affect child development (Olivennes *et al.*, 1993; Draper *et al.*, 1999; Jackson *et al.*, 2004; Pinborg *et al.*, 2005; Kapiteijn *et al.*, 2006; Sutcliffe and Ludwig, 2007; Griesinger *et al.*, 2008).

To study the potential effects of various components of the ART procedure separately, the Groningen ART-cohort study was initiated. In this study three prospectively recruited groups of children were included. The first group consisted of children born following a conventional, so-called 'controlled ovarian hyperstimulation'-IVF procedure (COH-IVF). The second group was born following IVF in the modified natural cycle (MNC-IVF). In this procedure, no ovarian hyperstimulation is performed (Rongières-Bertrand *et al.*, 1999) and, therefore, potential differences in outcome of COH-IVF and MNC-IVF children may be

attributed to the ovarian hyperstimulation. The third group consisted of naturally conceived (sub-NC) children born to subfertile couples. The comparison of MNC-IVF children and sub-NC children was used to study the net effect of the in vitro procedure. The differentiation of the effects of ovarian hyperstimulation and the in vitro procedure on neurodevelopmental outcome is a unique aspect of our study.

Previously, we reported on the neurodevelopmental outcome of children in the Groningen ART-cohort study at the ages of 2 weeks and 3 months (Middelburg *et al.*, 2009). At those ages, neurodevelopment of COH-IVF, MNC-IVF and sub-NC children was similar. However, continuation of follow-up is needed as children show a rapid expansion in functional repertoire during childhood. The demand for increasingly complex brain function may lead to the appearance of dysfunction when children grow older.

In the present study, we report on the neurological outcome of children in the Groningen-ART cohort at the ages of 4, 10 and 18 months. Standardised, age-specific and sensitive neurological assessments by blinded assessors allowed us to study potential minimal differences between the three conception groups. Primary outcome was neurological condition at 18 months expressed in terms of fluency of motor behaviour. This aspect of motor behaviour is easily affected by neurological dysfunction and is therefore a sensitive measure for minimal changes in neuromotor development (Huisman *et al.*, 1995). Secondary outcome measures were type and severity of minor neurological dysfunction (MND) at 4, 10 and 18 months and developmental trajectories from 4 until 18 months.

METHODOLOGY

Participants

For this longitudinal, prospective follow-up study, we recruited pregnant couples with a term date between March 2005 and December 2006 through the department of Reproductive Medicine of the University Medical Center Groningen (Middelburg *et al.*, 2009). All couples who achieved a singleton pregnancy following IVF or ICSI (either COH-IVF/ICSI or MNC-IVF/ICSI) were invited to participate. Details on treatment protocol and procedures in MNC-IVF have previously been described by Pelinck *et al.* (2007, 2008). Excluded from the study were couples with a pregnancy following cryopreservation or donation of oocytes or embryos. As a naturally conceived control cohort, we invited all couples who achieved a singleton pregnancy while on the waiting list for fertility evaluation or treatment during the study period (sub-NC). These couples had been subfertile for at least 1 year;

therefore, we expected that parental characteristics, such as parity and age, of this cohort would resemble the characteristics of IVF couples.

All couples were invited to participate during the third trimester of pregnancy. At the first appointment, approximately 2 weeks post-term, demographic information, such as parity, gestational age, birthweight, neonatal intensive-care unit admission, parental age and parental educational level, was collected on standardised charts. Information on time to pregnancy and the occurrence of vanishing twins were retrieved from fertility charts. The ethics committee of the University Medical Center Groningen approved the study design, and at least one of the parents provided written informed consent for participation of their infant in the study.

Neurological assessments

Follow-up consisted of standardised, age-specific neurological assessments at the ages of 4, 10 and 18 months post- term. Age-specific testing is necessary, since in infancy the nervous system shows many structural and functional changes.

At 4 and 10 months, we used the Touwen Infant Neurological Examination (TINE) to assess the neurological outcome (Touwen, 1976). In this assessment, neurological condition is summarized with the help of clusters of signs. The clusters are organized according to the functional, neurobehavioural subsystems of the nervous system used in clinical practice. Examples are fine motor function (reaching and grasping), gross motor function, brain stem function, visuomotor function and sensorimotor function. Each cluster can be scored as typical or deviant (criteria are reported in Hadders-Algra *et al.* (2009). Major neurological dysfunction means the presence of a distinct neurological syndrome, such as a hemisyndrome, irrespective of the number of deviant clusters. MND is scored when more than two clusters are deviant. Two forms of normal neurological condition are distinguished: normal-suboptimal, when one or two clusters are deviant and neurologically normal, when none of the clusters are deviant (Hadders-Algra *et al.*, 2009). The reliability of determining MND with TINE is good ($\kappa = 0.83$); construct validity of MND in infancy is good and predictive validity is moderate (Hadders-Algra *et al.*, 2009).

The neurological examination according to Hempel (1993) was used at 18 months. Basic principles of the TINE and Hempel are identical, but due to the substantial age-dependent changes in neuromotor behaviour, the assessments differ in contents of items and criteria for deviancy. In the Hempel assessment, the following clusters are scored as typical or deviant: fine-motor function, gross-motor function, posture and muscle tone, reflexes and visuomotor function (Hadders-Algra, 2003). Similar to the TINE classification, major neurological dysfunction implies the presence of a distinct neurological syndrome, such as cerebral palsy (CP). At this age

it is possible to make a distinction between two main categories of minor neurological dysfunction: complex MND and simple MND (Hadders-Algra, 2003). Complex MND is strongly related to preterm birth and perinatal adversities; it is the form of MND with clinical relevance due to its clear association with learning and behavioural disorders (Hadders-Algra, 2002; Batstra *et al.*, 2003). Complex MND is scored when two or more deviant clusters are present. Simple MND can be seen as a normal, but non-optimal form of brain function; it is scored when one cluster is deviant, i.e. the isolated presence of fine motor, gross motor or visuomotor dysfunction or mild dysregulation of posture and muscle tone. Neurologically normal implies the presence of no deviant clusters or only the presence of the cluster reflexes. The reliability of the Hempel examination is satisfactory (κ scores for various items: 0.62–1.00). Information on the predictive validity is lacking thus far (Hadders-Algra, 2005).

Our primary outcome measure was the fluency of motor behaviour at 18 months. The fluency score is a sub-score of the neurological optimality score (NOS) which is based on the Hempel examination. The items of the neurological examination have a predefined optimal range (Huisman *et al.*, 1995). The total number of items scored within the optimal range determines the NOS (range 0–58). It is important to realize that there is a conceptual difference between normality and optimality; the range for optimal behaviour is narrower than for normal behaviour (Prechtl, 1980). Due to this phenomenon, the NOS is able to evaluate subtle differences in neurological outcome. A sub-score of the NOS deals with fluency of motor behaviour (fluency score; range 0–13). Since subtle dysfunction of the nervous system is most easily expressed in a reduction of the fluency of movements, this measure is sensitive for minimal changes in neuromotor development.

At the ages of 4, 10 and 18 months children were assessed by KJM, who was blind to mode of conception. Parents were instructed not to reveal any information regarding conception method.

Statistical analysis

Power calculation of the longitudinal study is based on neurological outcome at the age of 18 months. For detection of at least half a standard deviation difference on the fluency subscore of the NOS (mean 9.5, standard deviation 1.7, (Huisman *et al.*, 1995), with 80% power, at least 64 children had to be included per group.

Mann–Whitney U-test or Student's t-test was used to compare the continuous variables, and chi-square test or Fisher's exact test to compare the categorical variables. The influence of ovarian hyperstimulation, the in vitro procedure or the combination of these two factors on neurological outcome was analysed using multiple regression analyses. The NOS, the fluency score and the occurrence of complex MND were used as dependent variables in, respectively,

linear and logistic regression analyses. The NOS and the fluency score had to be transformed, as residuals in the linear regression were non-normally distributed. The NOS was transformed into: –ln (59.5 – NOS), and the fluency sore was transformed into: – ln (14.5 – fluency score). We corrected for variables for which the groups differed at 5% significance level in the multivariate analyses. In addition, gestational age was entered in the multivariate analyses, since we know from literature that it is an important predictor for neurological outcome. We have used the results of multiple linear regression analysis to calculate confidence intervals (CI) for adjusted difference between the means of the three groups. To interpret these intervals on the original scale, we use the fact that the difference between means of two groups, A and B, on the transformed scale for the fluency score can be interpreted as the logarithm of the ratio (14.5 – medB)/(14.5 – medA), where medA and medB are medians on the original scale. Statistical analyses were performed using SPSS 14.0 for Windows. P-values of 5% or less were considered significant.

RESULTS

Participation and demographic characteristics

There were 89 children born following COH-IVF, 79 following MNC-IVF and 143 following a natural conception who were eligible for participation in the follow-up study. Parents of, respectively, 68 (76%), 57 (72%) and 90 (63%) children agreed to participate. Non-participants were similar to participants for gender, number of firstborn children, birthweight, prematurity-rate, neonatal intensive-care admission, parental educational level and time to pregnancy (results not presented). However, non-participating sub-NC mothers were significantly younger than participating sub-NC mothers (P = .03; data not presented).

Table I shows the demographic and perinatal characteristics of participating families. Overall, the groups were similar. Exceptions to this rule were the following: birthweight and gestational age were significantly higher and longer following sub-NC than following COH-IVF (P = .02, P = .02). Signs of fetal distress (denoted by meconium stained amniotic fluid, cardiotocographic signs or acidosis) were observed in 44% of children in the sub-NC group compared with 29% in the COH-IVF group (P = .054) and 28% in the MNC-IVF group (P = .046). Time to pregnancy was significantly shorter in the sub-NC group (median value 2.1 years) than in the COH-IVF group (4.1 years; P < .0005) and MNC-IVF group (3.8 years; P = .002). Eight children in the COH-IVF group were survivors of a vanishing twin compared with one in the MNC-IVF group (P = .04) and none in the sub-NC group (P = .001).

Table 1 - Demographic characteristics

Characteristics	COH-IVF[a] *n = 68*	MNC-IVF[a] *n = 57*	NC[a] *n = 90*
Gender: male, n (%)	36 (53%)	27 (47%)	46 (51%)
First born, n (%)	47 (69%)	38 (67%)	55 (61%)
Birth characteristics:			
Gestational age in weeks; median (range)	39.4 (33-42)*	40.1 (35-43)	40.0 (30-43)*
Preterm birth (< 37 weeks); n (%)	7 (10%)	6 (11%)	7 (8%)
Birth weight in grams; median (range)	3378 (1980-4700)*	3400 (2170-4680)	3565 (1150-4710)*
Low birth weight (< 2500 gram); n (%)	3 (4%)	4 (7%)	5 (6%)
Small for gestational age[b]; n (%)	0 (0%)	3 (5%)	2 (2%)
Forceps/ vacuum extraction, n (%)	6 (9%)	7 (12%)	11 (12%)
Caesarean section, n (%)	17 (25%)	8 (14%)	24 (27%)
Signs of fetal distress[c], n (%)	20 (29%)	16 (28%)*	40 (44%)*
Neonatal characteristicse:			
Apgar score 5 min < 7, n (%)	0 (0%)	0 (0%)	1 (1%)
Neonatal intensive-care admission, n (%)	1 (2%)	2 (4%)	7 (8%)
Parental characteristicse:			
Maternal age at conception; median (range)	32.5 (26-41)	32.8 (25-37)	33.2 (22-40)
Paternal age at conception; median (range)	35.7 (28-56)	34.4 (28-48)	35.4 (25-53)
Education level mother high[d], n (%)	22 (32%)	22 (39%)	41 (46%)
Education level father high[d], n (%)	29 (45%)	19 (34%)	33 (37%)
Fertility parameterse:			
Intracytoplasmic sperm injection; n (%)	43 (63%)	29 (51%)	na
Time to pregnancy in years; median (range)	4.1 (0-13)***	3.8 (0-13)**	2.1 (0-11)***/**
Vanishing twins, n (%)	8 (12%)*/**	1 (2%)*	0 (0%)**
Corrected age at examination:			
4 months; median in weeks (range)	18 (14-23)	18 (17-21)	18 (14-21)
10 months; median in weeks (range)	44 (42-56)	44 (39-48)	44 (41-51)
18 months; median in years (range)	1.5 (1.4-1.8)	1.5 (1.4-1.7)	1.5 (1.4-1.7)

Mann-Whitney U test and χ^2 or Fisher's exact test were used to compare groups; * P < .05, ** P < .017 (Bonferonni correction), *** P < .001.

[a] COH-IVF: infants born following controlled ovarian hyperstimulation IVF, MNC-IVF: modified natural cycle IVF, and NC: naturally conceived controls born to subfertile parents.

[b] Birth weight for gestational age is < -2 standard deviation scores compared to a Dutch reference population (Dutch reference tables, Perinatal Registration Netherlands).

[c] Signs of fetal distress denoted by meconium stained amniotic fluid and/or cardiotocographic signs and/or acidosis.

[d] University education or vocational colleges.

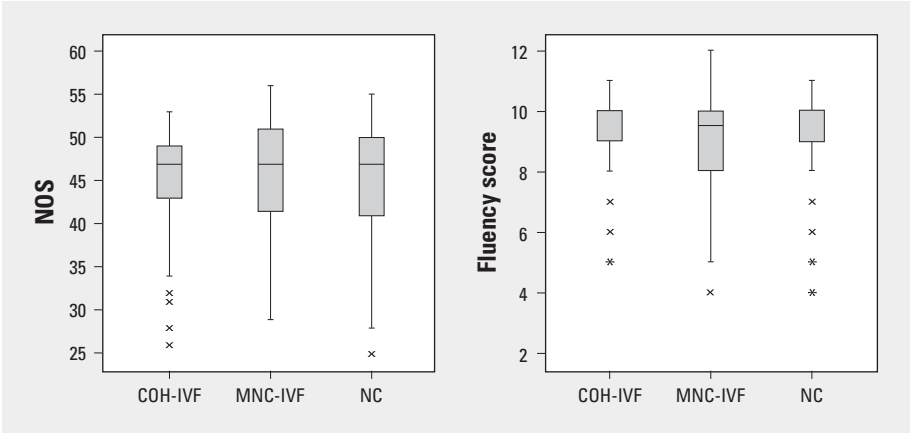
[e] Missing values for vaiabeles in the COH-IVF, MNC-IVF and NC group: < 3, <2, and <4, respectively.

Neurological optimality and fluency of movements at 18 months

Attrition at the 18-month assessment was minimal, two COH-IVF, one MNC-IVF and three sub-NC-children were lost to follow-up at 18 months. Five of these children were not assessed due to logistical reasons. One girl, born following MNC-IVF, died of a congenital heart disorder when she was 3 weeks old.

Figure 1 shows the distribution of the NOS and its fluency score for children in the three conception groups at 18 months. The median score of the NOS was 47 in all conception groups. The median value of the fluency score was 10 in the COH-IVF group, 9.5 in MNC-IVF and 9 in sub-NC children, these differences were statistically non-significant. Multiple linear regression confirmed that neither the ovarian hyperstimulation (COH-IVF versus MNC-IVF), nor the in vitro procedure (MNC-IVF versus sub-NC) nor a combination of these two factors (COH-IVF versus sub-NC) influenced the NOS or the fluency score (table II). Transforming the CIs for the differences between groups (table II) back to the original scale results in the following interpretation: assuming that the corrected median fluency score is 9 in the sub-NC group, the CIs for corrected medians for the MNC-IVF and COH-IVF groups are (8.3–9.5) and (8.6–9.7) respectively; assuming the score is 9.5 for the MNC-IVF group, the CI for the COH-IVF group median is (9.2– 10.2). For NOS, assuming median score of 47 for the sub-NC group results in CIs (45.9 –49.4) and (44.9–48.7) for medians in the MNC-IVF and COH-IVF groups, respectively; assuming the score is 47 for the MNC-IVF group, the CI for the COH-IVF group is (44.0–48.0).

Figure 1 - distribution of the NOS and fluency score at 18 months



COH-IVF: infants born following controlled ovarian hyperstimulation IVF, MNC-IVF: modified natural cycle IVF, and sub-NC: naturally conceived controls born to subfertile parents. Differences between groups in the NOS and fluency score were non-significant.

Minor neurological dysfunction from 4 until 18 months

Neurological outcome at various ages is presented in table III. None of the children showed major neurological dysfunction. At the age of 4 and 10 months, the rate of children classified as normal, normal-suboptimal or MND was similar in the COH-IVF, MNC-IVF and sub-NC groups. Also at 18 months, we observed similar rates of

Table II - Multivariate regression analyses of contribution of IVF-method on outcome measures at 18 months.

Outcome measure: Neurological optimality score (-ln (59.5 - NOS))

compared groups	adjusted median difference [95% CI][e]	p-value
COH-IVF vs. MNC-IVF	-0.066 [-0.226, 0.094][b]	0.415
MNC-IVF vs. NC	0.067 [-0.084, 0.218][c]	0.385
COH-IVF vs. NC	-0.004 [-0.155, 0.146][d]	0.957

Outcome measure: Fluency score (-ln (14.5 - Fluency score))

compared groups	adjusted median difference [95% CI][e]	p-value
COH-IVF vs. MNC-IVF	0.049 [-0.055, 0.152][b]	0.354
MNC-IVF vs. NC	-0.010 [-0.116, 0.095][c]	0.846
COH-IVF vs. NC	0.038 [-0.065, 0.142][d]	0.464

Outcome measure: Complex MND[a]

compared groups	adjusted odds ratio [95% CI][f]	p-value
COH-IVF vs. MNC-IVF	1.30 [0.38, 4.42][b]	0.670
MNC-IVF vs. NC	1.44 [0.40, 5.25][c]	0.576
COH-IVF vs. NC	1.85 [0.55, 6.24][d]	0.319

COH-IVF: infants born following controlled ovarian hyperstimulation IVF, MNC-IVF: modified natural cycle IVF, and NC: naturally conceived controls born to subfertile parents.

[a] MND: Minor Neurological Dysfunction
[b] Adjusted for gestational age and vanishing twins.
[c] Adjusted for gestational age, signs of fetal distress and time to pregnancy.
[d] Adjusted for gestational age, birth weight, vanishing twins and time to pregnancy.
[e] Multivariate linear regression was used for NOS and fluency score.
[f] Multivariate logistic regression was used for Complex MND.

children presenting with a normal neurological outcome, simple MND or complex MND in the three groups. At all ages, specific clusters of dysfunction occurred equally frequent in the three groups. An exception was sensorimotor dysfunction at the age of 10 months; this was observed in 38% of children born following COH-IVF compared with 18% of MNC-IVF children (P = .015) and 27% of the sub-NC children (only data on specific clusters at 18 months are presented). Logistic regression analysis with correction for confounders confirmed that conception method did not explain the presence of complex MND. Table II shows the adjusted odds ratios for the effects of ovarian hyperstimulation, the in vitro procedure or a combination of these two factors on complex MND at the age of 18 months. Results of logistic regression analysis at the ages of 4 and 10 months were similar to those at 18 months (data not presented).

TABLE III - NEUROLOGICAL CLASSIFICATION AND CLUSTERS OF DYSFUNCTION.

Outcome measure	COH-IVF[a] *n = 68*	MNC-IVF[a] *n = 57*	NC[a] *n = 90*
Neurological outcome at 4 months			
Normal	35 (52%)	25 (45%)	38 (44%)
Normal-suboptimal	28 (41%)	29 (52%)	45 (52%)
Minor Neurological Dysfunction	5 (7%)	2 (4%)	3 (4%)
Total number of children tested	68	56	86
Neurological outcome at 10 months			
Normal	36 (54%)	36 (64%)	55 (62%)
Normal-suboptimal	27 (41%)	20 (36%)	33 (37%)
Minor Neurological Dysfunction	3 (4%)	0 (0%)	1 (1%)
Total number of children tested	66	56	89
Neurological outcome at 18 months			
Normal	44 (67%)	40 (71%)	69 (79%)
Simple MND[b]	15 (23%)	11 (20%)	12 (14%)
Complex MND[b]	7 (11%)	5 (9%)	6 (7%)
Total number of children tested	66	56	87
Clusters of dysfunction at 18 months			
Fine motor dysfunction	0 (0%)	1 (2%)	2 (2%)
Gross motor dysfunction	21 (32%)	14 (25%)	17 (20%)
Posture and muscle tone dysfunction	2 (3%)	3 (5%)	2 (2%)
Dysfunctional reflexes	15 (23%)	11(20%)	23 (27%)
Visuomotor dysfunction	1 (2%)	0 (0%)	0 (0%)

No significant differences ($p > 0.05$), Mann-Whitney, χ^2 or Fisher's Exact test.

[a] COH-IVF: infants born following controlled ovarian hyperstimulation IVF, MNC-IVF: modified natural cycle IVF, and NC: naturally conceived controls born to subfertile parents.

[b] MND: Minor Neurological Dysfunction

Table IV shows the different developmental trajectories observed in the three groups. At 4 and 10 months, we dichotomized outcome into normal (normal and normal-suboptimal) and MND, and at 18 months into normal (normal and simple MND) and complex MND. The large majority of children showed a consistent normal developmental trajectory. Rates of children with a consistent normal neurological condition from 4 until 18 months were similar in the three groups, i.e. 85% of children born following COH-IVF, 88% of MNC-IVF children and 92% of sub-NC children. Neurological outcome improved with age in three (5%) COH-IVF children, two (4%) MNC-IVF children and one (1%) sub-NC child. It deteriorated in six (9%), five (9%) and six (7%) children, respectively. The rates of improvement and deterioration were not significantly different between the groups. Only one child who was born following COH-IVF consistently showed MND or complex MND.

Table IV - Developmental trajectories from 4 until 18 months.

Developmental Traject[b]	COH-IVF[a] *n = 66*	MNC-IVF[a] *n = 56*	NC[a] *n = 84*
Neurological classification at 4, 10, and 18 months			
normal, normal, normal	56 (85%)	49 (88%)	77 (92%)
MND, normal, normal[c]	2 (3%)	2 (4%)	1 (1%)
normal, MND, normal[c]	1 (2%)	0 (0%)	0 (0%)
MND, MND, normal[c]	0 (0%)	0 (0%)	0 (0%)
normal, normal, C-MND[d]	3 (4%)	5 (9%)	3 (4%)
MND, normal, C-MND[d]	2 (3%)	0 (0%)	2 (2%)
normal, MND, C-MND[d]	1 (2%)	0 (0%)	1 (1%)
MND, MND, C-MND	1 (2%)	0 (0%)	0 (0%)

No significant differences (p >0.05), Chi-square or Fisher's Exact test.

[a] COH-IVF: infants born following controlled ovarian hyperstimulation IVF, MNC-IVF: modified natural cycle IVF, and NC: naturally conceived controls born to subfertile parents.

[b] Only children who were assessed at all ages were used in this analysis. MND: minor neurological dysfunction, C-MND: complex - minor neurological dysfunction.

[c] Improvement in developmental trajectory with age.

[d] Deterioration in developmental trajectory with age.

DISCUSSION

The present study that used highly sensitive measures, found no effects of ovarian hyperstimulation or the in vitro procedure itself on neurological outcome in children aged 4–18 months. Fluency of movements, neurological optimality and the occurrence of complex MND were similar between children born following COH-IVF, MNC-IVF and their naturally conceived peers born to subfertile parents.

The findings in the present study support those of a previous stage of the study, when children were 2 weeks to 3 months old (Middelburg *et al.*, 2009). This means that we found no significant differences in neurodevelopmental outcome between COH-IVF, MNC-IVF and sub-NC children at five ages ranging from the neonatal period to toddler age. The consistency of the finding is especially reassuring, as generally, multiple neurological assessments have a better predictive validity than single assessments (Hadders-Algra *et al.*, 2009). At the ages of 2 weeks and 3 months, we additionally studied the effect of subfertility itself on early neurodevelopmental outcome by comparing children of subfertile parents (the COH-IVF, MNC-IVF and sub-NC group together) to children in a reference population (Middelburg *et al.*, 2009). Results suggested that rather than the ART procedures, subfertility itself was associated with less-optimal neurodevelopmental outcome (Middelburg *et al.*, 2009). Unfortunately, the design of the study on

developmental outcome in the general population did, however, not allow for a detailed longitudinal assessment of all infants (Bouwstra *et al.*, 2009).

The longitudinal design of our study allowed us to analyse the developmental trajectories of the children in the three conception groups. A large majority of the children showed a normal developmental trajectory up to 18 months. Rates of improvement or deterioration of neurological outcome from 4 to 18 months were similar in the three conception groups, indicating that IVF children neither have to catch-up form early deviancies in development, nor do they grow into minor neurological dysfunction up to the age of 18 months.

The control group in this study was composed of children born to subfertile parents. With the inclusion of this group, we aimed to compose a control group that resembled the IVF groups in demographic characteristics, so that the effect of potential confounders, such as parity and maternal age was minimized. Nevertheless, time to pregnancy was significantly shorter for couples who conceived naturally. Possibly, this indicates that the sub-NC couples were less subfertile. Gestational age and birthweight of children in the COH-IVF group were lower than those of sub-NC children. Whereas, in contrast, signs of fetal distress, NICU admission and Caesarean section were more frequently observed in sub-NC children. The increased risk of adverse perinatal outcome in our sub-NC group may have been a chance finding, but may also have been the result of differences in obstetrical care between ART and naturally conceived pregnancies. We corrected for the differences in perinatal outcome by means of multivariate statistics as our research question was whether ovarian hyperstimulation or the in vitro procedure affected neurological outcome, given the potential effect of assisted conception on perinatal outcome. It is, however, also arguable not to correct for these factors, since they might be mediating factors on the causal pathway from assisted conception to neurological outcome. A different approach would, however, not have essentially changed our results. The univariate statistics were not significant and P-values changed only marginally after correction for confounders.

The prospective design of this study, in which couples were invited during pregnancy, reduced the chance of selection bias based on the child's health or development. In the initial phase of the study, 63–76% of eligible children were included. Since characteristics of participants and non-participants were similar, we assume that the children we included are a representative sample. Given the intensity of the study (five assessments in 18 months), the follow-up percentage up to 18 months (97–98%) is high. Except for the child who died of a congenital heart disorder, there was no reason to expect dropout to be selective.

Blinding of the assessor to mode of conception was a strength of our study. Recently, it was questioned whether blinding in ART follow-up studies is adequate,

since factors such as a child's singleton status, parental age and parental behaviour may provide clues about the child's mode of conception (Ludwig *et al.*, 2009). In the present study, comparability of study and control families was enlarged by the fact that control parents had also experienced subfertility. Importantly, the number of firstborn children was similar in the ART and control groups. Therefore, the likelihood that the assessor would be able to guess conception mode was further reduced.

A limitation of our study was that sample size in the MNC-IVF group turned out to be slightly smaller than the 64 children needed to detect half a standard deviation difference in the fluency score. Partly, this was compensated for by larger groups of COH-IVF and sub-NC children. Further, the lack of a trend for worse or better outcome in one of the groups makes it unlikely that a slightly larger sample size would have led to a significant difference between groups.

Another limitation of our study design is that the minimal medication used in MNC-IVF may cause an overestimation of the effect of the in vitro procedure or an underestimation of the effect of ovarian hyperstimulation. This minor confounding of MNC was not so much a problem for interpretation of the results of this study, since we found no effect for both procedures.

Previous studies on neuromotor development in IVF children have often used relatively gross measures of neurodevelopmental outcome, such as the Bayley or Griffiths scales, which were not designed to study neurological outcome in a detailed sense. With these instruments, potentially subtle differences between groups could have remained undetected. On the level of an individual child, these subtle differences in outcome might seem of little clinical relevance. However, on a population level such differences start to matter. For instance, a three point reduction in intelligence quotient (IQ) may not have direct consequences for an individual child. But, when IQ in 4% of children decreases with three points, this may have serious consequences for society at large, on long term. Furthermore, children with scores at the lower edge of the normal range may cross borders to scores beneath the normal range (Knoester *et al.*, 2008). It is also important to realize that in statistical analyses, we are able to correct for confounding factors between ART and non-ART children (e.g. gestational age, maternal age and parity), but the crude differences in outcome remain in the population and have their consequences for society.

The findings of our study in IVF children up to 18 months are reassuring. It should, however, be realised that subtle neurodevelopmental disorders may emerge when children grow older. Therefore, continuation of follow-up in older children is needed. To our knowledge, only two studies have reported on minor deviancies in neurological outcome of pre-school or school-age ART children

(Middelburg *et al.*, 2008). Recently, Knoester *et al.* (2007) observed in a thoroughly matched, assessor-blinded cohort study a similar prevalence of MND in 5–8-year-old IVF and ICSI children. In the same study, a higher crude prevalence of MND was observed in ICSI children compared with naturally conceived children; however, after adjustment for confounders (most importantly parity), this difference was no longer statistically significant. A striking finding of Knoester's study was the high rate of MND in the study as well as the control group (simple + complex MND; ICSI: 66%, IVF: 61% and naturally conceived 51%) (Knoester *et al.*, 2007). The high rate of MND in their naturally conceived group may be a sign of selection bias; parents with worries concerning their child's motor performance may have been keen to volunteer for the naturally conceived control group. Theoretically, such confounding could have concealed a difference in occurrence of MND between ART and naturally conceived children. Another study, performed by Belva *et al.* (2007) found no substantial differences in neurological outcome between 8-year-old ICSI and naturally conceived children. Unfortunately, the latter study analysed the items of the Touwen neurological examination separately, and refrained from summarizing results in dysfunctional clusters, so that important information on the prevalence of MND was lost. This means that additional studies focusing on subtle neurological dysfunction beyond infancy in ART children are highly warranted.

In conclusion, in this longitudinal, prospective, assessor-blinded cohort study, we observed similar neurological outcomes in children born following COH-IVF, MNC-IVF and children of subfertile couples, up to the age of 18 months. In order to be able to detect subtle differences, we studied neurological outcome with detailed and standardised neurological assessments. The absence of differences between the groups suggests that neither ovarian hyperstimulation, nor the in vitro procedure affect neurological outcome in early childhood. Long-term follow-up, in large groups of children, focusing on subtle neurological dysfunction is still needed to confirm our findings.

Part II
The Groningen
ART cohort study

5

The effects of ovarian hyperstimulation and the IVF laboratory procedures on neurological condition at 2 years

P. Schendelaar
K.J. Middelburg
A.F. Bos
M.J. Heineman
M. Jongbloed-Pereboom
M. Hadders-Algra

Human Reproduction 2011; 26: 703-712.

ABSTRACT

Background: Up to 4% of children are born following assisted reproduction techniques (ART) yet relatively little is known on neurodevelopmental outcome of these children after 18 months of age. Only a limited number of long-term follow-up studies with adequate methodological quality have been reported. Our aim was to evaluate the effects of ovarian hyperstimulation, IVF laboratory procedures and a history of subfertility on neurological condition at 2 years.
Methods: Singletons born after controlled ovarian hyperstimulation IVF (COH-IVF, n= 66), modified natural cycle IVF (MNC-IVF, n= 56), natural conception in subfertile couples (sub-NC, n= 87) and in fertile couples (reference group, n= 101) were assessed (using Hempel approach) by neurological examination at 2 years of age. This resulted in a neurological optimality score (NOS), a fluency score and the prevalence of minor neurological dysfunction (MND). Primary outcome was the fluency score, as fluency of movements is easily affected by subtle dysfunction of the nervous system.
Results: Fluency score, NOS and prevalence of MND were similar in COH-IVF, MNC-IVF and sub-NC children. However, the fluency score ($P < .01$) and NOS ($P < .001$) of the three subfertile groups were higher, and the prevalence of MND was lower ($P = .045$), than those in the reference group.
Conclusions: Neurological condition of 2 year olds born after ART is similar to that of children of subfertile couples conceived naturally. Moreover, subfertility does not seem to be associated with a worse neurological outcome. These findings are reassuring, but we have to keep in mind that subtle neurodevelopmental disorders may emerge as children grow older.

INTRODUCTION

Worldwide, already 3 million children have been born following assisted reproduction techniques (ART) (Adamson *et al.*, 2006). This means that even subtle adverse effects of ART on neurodevelopmental outcome will have consequences for society at large.

A recent systematic review indicated that ART does not seem to affect short-term neurodevelopmental outcome (Middelburg *et al.*, 2008). This review also revealed that surprisingly little is known on long-term outcome of ART owing to the relative paucity of long-term follow-up studies of good methodological quality. Only seven good quality studies of children older than 18 months were included in the review (Morin *et al.*, 1989; Brandes *et al.*, 1992; Levy-Shiff *et al.*, 1998; Agarwal *et al.*, 2005; Wikstrand *et al.*, 2006; Knoester *et al.*, 2007; Knoester *et al.*, 2008). Since that time, only one other good quality study has been published (Ludwig *et al.*, 2009a). In general, the follow-up studies indicated that neurodevelopmental outcome of ART children is similar to that of children conceived naturally. However, some studies reported inconsistent results. For instance, Knoester *et al.* (2008) indicated that cognitive function of ICSI children was worse than that of children born following IVF and natural conception. Levy-Shiff *et al.* (1998) reported an excess of behavioural problems in IVF boys, whereas Morin *et al.* (1989) reported more vocalization and higher energy levels in IVF children compared with naturally conceived peers. Neurodevelopmental outcome after ART has also been assessed by means of register-based studies (see review of Middelburg *et al.*, 2008). These studies suggested that IVF/ICSI per se does not increase the risk for severe cognitive impairment or neurological handicap, such as cerebral palsy. A disadvantage of register-based studies, however, is that minor effects of ART on neurological condition cannot be evaluated.

Altogether, this means that we are currently still not well informed about potentially long-term sequelae of ART on neurological outcome. The finding that ART is associated with perinatal complications such as preterm birth and low birthweight (Schieve *et al.*, 2002; Helmerhorst *et al.*, 2004; Jackson *et al.*, 2004), suggests that ART might be associated with a minor negative impact on more advanced neurodevelopmental functions.

In theory, different components of ART may affect development. Hypothetical points of concern are ovarian hyperstimulation (Griesinger *et al.*, 2008), the impact of the ART procedure itself (Olivennes *et al.*, 1993) and consequences of vanishing twins (Pinborg *et al.*, 2005). In addition, background factors associated with ART, such as a history of subfertility resulting from tuba pathology, andrological factors or increased maternal age (Salem Yaniv *et al.*,

2010), may contribute to worse neurodevelopmental outcome in ART singletons (Draper *et al.*, 1999; Sutcliffe and Ludwig, 2007).

To examine potential effects of the various components of ART, the Groningen ART cohort was composed (figure 1; Middelburg *et al.*, 2009a,b). The cohort consists of three groups. The first group consisted of children born after conventional controlled ovarian hyperstimulation IVF (COH-IVF), the second of children born after IVF in the modified natural cycle (MNC-IVF) in which medication use was minimal (Nargund *et al.*, 2007) and the third group consisted of a control group of naturally conceived children born to subfertile couples (Sub-NC). Possible differences in outcome of COH-IVF and MNC-IVF children may be attributed to ovarian hyperstimulation, whereas potential differences in MNC-IVF and Sub-NC children may largely be attributed to the ART procedure.

We previously reported that neurological outcome of children of the Groningen ART cohort was similar at 2 weeks and 3 months of age (Middelburg *et al.*, 2009a) and up to and including 18 months (Middelburg *et al.*, 2009b). However, neurological condition at 3 months of the Groningen ART cohort was less favourable than that of a reference group of the general population, suggesting that neither ovarian hyperstimulation, nor the in vitro procedure, but rather factors associated with subfertility affect early neurodevelopmental outcome (Middelburg *et al.*, 2009a).

The primary aim of this paper is to assess the effect of ovarian hyperstimulation and the IVF laboratory procedure on neurological outcome at 2 years. To this end, all children of the prospective Groningen ART cohort were re-assessed using the standardised, precise and age-specific neurological assessment according to Hempel (1993). Age-specific testing is necessary, since many functional and structural changes of the brain occur during childhood. Primary outcome was the neurological condition expressed in terms of fluency of motor behaviour. This is a sensitive measure to detect subtle changes in neuromotor development, since minor dysfunction of the nervous system already results in a reduction of fluency of motion (Huisman *et al.*, 1995; Hadders-Algra *et al.*, 2004). Secondary outcome parameters were the neurological optimality score (NOS) and type and severity of minor neurological dysfunction (MND). Specific attention was paid to relationships between causes of subfertility and neurological outcome. The secondary aim of the study was to compare neurological outcome of the Groningen ART study with that of a newly recruited retrospective reference group of 2 year olds born to fertile couples.

METHODOLOGY

Participants

Pregnant subfertile couples with a term date between March 2005 and December 2006 were recruited at the department of Reproductive Medicine of the University Medical Center Groningen (Middelburg *et al.*, 2009a). All couples who achieved a singleton pregnancy following IVF or ICSI were invited to participate, resulting in groups of children born after COH-IVF and MNC-IVF. Couples with a pregnancy after treatment with cryopreserved or donated oocytes or embryos were excluded. The third group formed was Sub-NC, which comprised couples that had tried to conceive for at least 1 year, and finally conceived naturally while waiting for fertility evaluation or treatment. We restricted our analysis to singletons, as -in contrast to COH-IVF- MNC-IVF rarely results in multiple gestation, and being a member of a multiple is associated with an increased risk for developmental problems (Pinborg *et al.*, 2003).

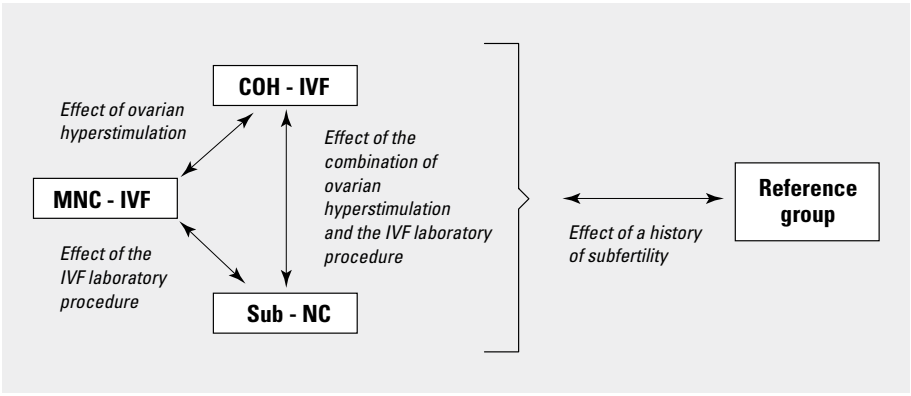
For the present study, a new retrospective reference group was recruited between February and October 2009 at six child welfare clinics in and around Groningen. All parents of 2-year-old children that visited the child welfare clinic for routine general health care were invited to participate. Children of couples who had attempted to achieve pregnancy for more than 1 year or achieved pregnancy by any form of assisted conception were excluded.

Prenatal, perinatal and demographic information had been gathered by use of standardised charts during the first follow-up assessment, ~2 weeks post-term (Middelburg *et al.*, 2009a). When information was incomplete or complications during pregnancy or birth had occurred, midwifes and gynaecologists were asked for details. Detailed information on the causes and treatment of infertility was retrieved from medical records. The Medical Ethical Commission of the University Medical Center Groningen approved the study design. At least one of the children's parents provided written informed consent for participation of their child in the study.

Neurological assessment

All children were assessed around the time of their second birthday. Neurological assessment was carried out according to Hempel: this assessment is a standardised neurological examination to assess MND at preschool age (Hempel, 1993). Five domains of function were assessed which can be scored as typical or deviant: fine motor function, gross motor function, posture and muscle tone, reflexes and visuomotor function (denoting function of the visual system and eye movements; Hadders-Algra *et al.*, 2004).

Figure 1 - The Groningen ART cohort study.



The effects of ovarian hyperstimulation, the IVF laboratory procedure, the combination of both and a history of subfertility are studied by four different comparisons. The groups controlled ovarian hyperstimulation IVF (COH-IVF), modified natural cycle IVF (MNC-IVF) and natural conception in subfertile couples (Sub-NC) were pooled to form the subfertile group.

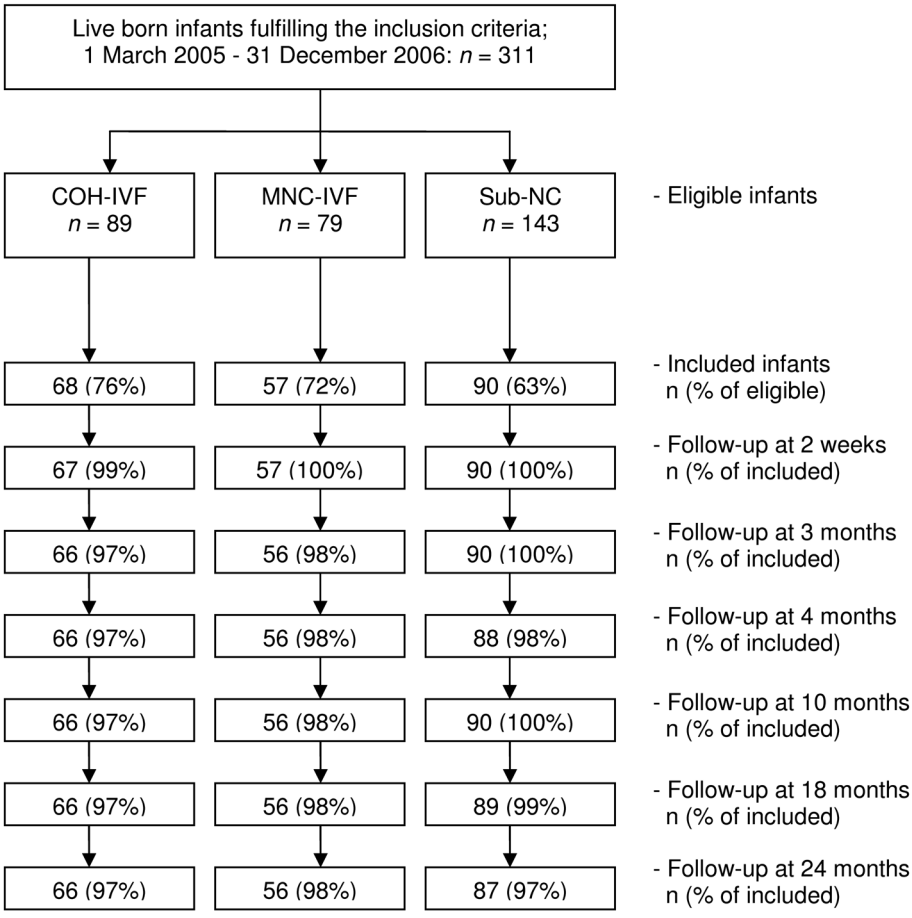
Children were classified as neurologically normal, simple MND, complex MND or neurologically abnormal. Simple MND denotes the presence of one deviant domain, and is regarded as a non-optimal, yet normal form of brain function (Hadders-Algra, 2002). Complex MND means the presence of more than one domain of dysfunction and represents the clinically relevant form, as it is associated with preterm birth and perinatal adversities, and behavioural and learning disorders (Hadders-Algra, 2002; Batstra *et al.*, 2003). Neurologically abnormal implies the presence of a distinct neurological syndrome, such as cerebral palsy. Neurologically normal implies the absence of neurological dysfunction and is scored when no domains are deviant or only the domain of reflexes is deviant.

In addition, outcome of the Hempel assessment was expressed in a NOS. The NOS consists of 58 items, for which an optimal condition is defined (range: 0–58). The sum of the items which fulfil the criteria for optimality forms the score. Higher scores represent better performance. It is important to realize that there is a conceptual difference between normality and optimality, since the range for optimal behaviour is narrower than for normal behaviour (Prechtl, 1980). This implies that both scores are suitable tools to evaluate subtle differences in neurological outcome. The fluency score (range: 0–13), a sub-score of the NOS, deals with the fluency of motor behaviour. Since subtle dysfunction of the nervous system is most easily expressed in a reduction of the fluency of movements, this measure is the most sensitive one to detect minimal changes in neuromotor development (Huisman *et al.*, 1995; Hadders-Algra *et al.*, 2004). The Hempel assessment has a satisfactory inter-rater reliability [$\kappa = 0.62–1.00$ (mean 0.93)] and

a good construct validity (Hadders-Algra, 2005). Note that in the results section neurological outcome is first presented in the clinical terms of MND. Thereafter, the results of the primary outcome parameter, the fluency score and the other additional outcome parameter, the NOS, are described.

M.J.P. and K.J.M., who assessed the children of the subfertile groups, were unaware of the mode of conception. Parents were asked not to reveal information about the conception method. However, the reference group was recruited separately from the subfertile groups, therefore it was impossible to keep the assessor of this group (M.J.P.) blind.

Figure 2 - Flow diagram of the Groningen ART cohort.



Statistical analysis

Power calculation was based on neurological outcome at 18 months. To detect at least half a SD difference in the fluency score (mean 9.5, SD 1.7) (Huisman *et al*., 1995), with 80% power at least 64 children had to be included per group (Middelburg *et al*., 2009b).

χ^2 test, Fisher's exact test, Mann–Whitney U-test or Student's t-test were used to test differences between groups. Potential associations between group status and neurological outcome were analysed using regression analyses. Linear regression analysis was used for the fluency score and NOS: to this end, both scores were normalised using –log (14.5–fluency score) and –ln (59.5–NOS), respectively. Logistic regression analysis was applied in the analysis of group effects on the occurrence of MND. A similar set of analyses was applied to assess the effect of specific causes of subfertility on neurological outcome.

In the multivariate analyses, background variables were included for which the groups differed at 5% significance level, with the exception of the subfertility causes. In addition, gestational age was entered in the multivariate analyses on an a priori basis, as gestational age is tightly linked to neurological outcome.

The results of the multiple linear regression analyses were used to calculate confidence intervals (CIs) for adjusted differences between the means of the groups. To interpret these on the original scale, we use the fact that the difference between means of two groups, A and B, on the transformed scale for the fluency score can be interpreted as the common logarithm of the ratio (14.5 – medB)/(14.5 – medA), where medA and medB are medians on the original scale.

Statistical analyses were performed using the Statistical Package for the Social Sciences 15.0 for Windows. Bonferroni corrections were applied in the analyses of the outcome parameters, but not in the analyses of patient characteristics. P-values of 5% or less were considered significant.

RESULTS

Participation and demographic characteristics

Eighty-nine children born after COH-IVF, 79 MNC-IVF children and 143 Sub-NC children met the inclusion criteria during the prenatal period: parents of 68 (76%), 57 (72%) and 90 (63%) children, respectively, agreed to participate in neurodevelopmental follow-up (figure 2; Middelburg *et al*., 2009a). Overall, obstetrical, neonatal and social characteristics of participants and non-participants were similar, except for maternal age (Middelburg *et al*., 2009a).

Table 1 - Characteristics of parents and infants of the Groningen ART cohort

Characteristics	COH-IVF	MNC-IVF	Sub-NC
	n = 66	*n = 56*	*n = 87*
Male gender; n (%)	36 (55)	27 (48)	45 (52)
First born, n (%)	45 (68)	38 (68)	53 (61)
Birth characteristics:			
Gestational age in weeks; median (range)	39.4 (33.4-42.3)*	40.1 (34.6-42.6)	40.0 (30.1-42.7)*
Preterm birth (< 37 weeks); n (%)	7 (11)	6 (11)	7 (8)
Birth weight in grams; mean (sd)	3396 (550)	3400 (576)	3547 (587)
Low birth weight (< 2500 gram), n (%)	3 (5)	4 (7)	5 (6)
Small for gestational age[a], n (%)	0 (0)	3 (5)	2 (2)
Caesarean section, n (%)	16 (24)	8 (14)	23 (26)
Signs of fetal distress[b], n (%)	19 (29)	16 (29)	38 (44)
Neonatal characteristics:			
Apgar score 5 min < 7[c], n (%)	0 (0)	0 (0)	1 (1)
Neonatal intensive care admission, n (%)	1 (2)	2 (4)	6 (7)
Breasfed for > 6 weeks[c], n (%)	30 (48)	26 (46)	42 (49)
Parental characteristics:			
Maternal age at conception in years; median (range)	32.9 (26.3-40.9)	32.8 (25.3-37.5)	33.3 (22.2-40.3)
Paternal age at conception in years[c]; median (range)	35.7 (27.5-56.1)*	34.4 (28.3-47.8)*	35.4 (25.5-52.6)
Smoking during pregnancy, n (%)	7 (11)	7 (13)	9 (10)
Parental socioeconomic status:			
Education level mother (high[d]), n (%)	22 (33)	21 (38)	40 (46)
Education level father (high[d])[c], n (%)	29 (46)	18 (32)	32 (47)
ICSI performed; n (%)	42 (64)	28 (50)	n.a.
Vanishing twins, n (%)	8 (12)*/***	1 (2)*	0 (0)***
Time to pregnancy in years[c]; median (range)	4.0 (0.1-13.3)***	3.8 (0.1-13.2)**	2.1 (0.1-11.3)***/**
Type of infertility (primary), n (%)	35 (53)	33 (59)	45 (52)
Subfertility causes[e]			
Tuba pathology, n (%)	8 (15)	9 (17)	5 (6)
Male factor, n (%)	28 (52)**	31 (57)***	19 (24)**/***
Other causes, n (%)	9 (17)	5 (9)	9 (11)
Unknown cause, n (%)	8 (15)***	9 (17)***	48 (59)***/***
Corrected age at examination at 2 years of age (in months)	24.9 (23.3-30.2)	24.9 (13.8-27.9)	25.0 (23.2-28.9)

Note: Mann-Whitney U tests or Student's t-test and Chi-square tests or Fisher's exact tests were used to compare between groups; * P < .05; ** P < .01, *** P < .001.

[a] Birth weight for gestational age is < -2 standard deviations compared with the Dutch reference population (Dutch reference tables, Perinatal Registration Netherlands).

[b] Signs of fetal distress denoted by meconium stained amniotic fluid and/or cardiotocographic signs and/or acidosis.

[c] Missing data in three groups: Apgar n=5, breastfed n=5, paternal age n=4, education level father n=4, time to pregnancy n=1. Note that in the percentages missing values have been taken into account.

[d] University education or vocational colleges.

[e] Couples may have more than 1 cause of subfertility, therefore totals may exceed 100%.

Six children were lost to follow-up at the assessment at 2 years. Two COH-IVF children and three Sub-NC children did not participate for logistical reasons. One MNC-IVF girl had a congenital heart disorder and died at 3 weeks of age.

Overall, demographic characteristics of the three groups were similar. The differences found were the following: gestational age was shorter in COH-IVF children than in Sub-NC children ($P = .012$). COH-IVF children were more often survivors of a vanishing twin than MNC-IVF and Sub-NC children ($P = .038$ and $P < .001$, respectively). Mothers of COH-IVF and MNC-IVF children needed more time to get pregnant than mothers of Sub-NC children ($P < .001$ and $P = .003$, respectively). Finally, fathers of MNC-IVF children were younger than those of COH-IVF children ($P = .037$; table I).

Table II - Neurological classification and domains of dysfunction.

a. ART cohort and reference group

	COH – IVF	MNC – IVF	Sub – NC	Subfertile group[a]	Reference group[b]
	n = 66	*n = 56*	*n = 87*	*n = 209*	*n = 101*
Neurological outcome					
Normal	58 (88%)	52 (93%)	83 (95%)	193 (92%)	88 (87%)
Simple MND[c]	5 (8%)	3 (5%)	2 (2%)	10 (5%)	10 (10%)
Complex MND[c]	3 (5%)	1 (2%)	2 (2%)	6 (3%)	3 (3%)
Domain of dysfunction					
Fine motor dysfunction	– (0)	1 (2%)	– (0)	1 (1%)	– (0)
Gross motor dysfunction	4 (6%)	1 (2%)	3 (3%)	8 (4%)	3 (3%)
Posture and muscle tone dysfunction	4 (6%)	2 (4%)	1 (1%)	7 (3%)*	12 (11%)*
Dysfunctional reflexes	11 (17%)	9 (16%)	12 (14%)	32 (15%)	22 (22%)
Visuomotor dysfunction	1 (2%)	– (0)	– (0)	1 (1%)	– (0)

b. Subfertility causes[d]

	Tuba pathology	Male factor	Other causes[e]	Unknown cause
	n = 29	*n = 94*	*n = 40*	*n = 65*
Neurological outcome				
Normal	27 (93)	85 (90)	37 (93)	60 (92)
Simple MND[c]	2 (7)	6 (6)	2 (5)	2 (3)
Complex MND[c]	– (0)	3 (3)	1 (3)	3 (5)

Note: χ^2 test or Fisher's Exact test were used to compare groups. * $P < .01$.

[a] Group COH-IVF, MNC-IVF and Sub-NC were pooled to form the subfertile group.

[b] Naturally conceived children born to fertile parents.

[c] MND = minor neurological dysfunction.

[d] Couples may have more than 1 cause of subfertility, therefore totals may exceed 100%.

[e] Other known causes of subertility: endometriosis, cervical factor, hormonal cause, lesbian couple.

Neurological condition of children of the Groningen ART cohort

None of the children of the COH-IVF, MNC-IVF and Sub-NC groups showed a definitely abnormal neurological condition, and none had an ICD-10 neurodevelopmental diagnosis. Children of the COH-IVF group tended to show both simple and complex MND more often than MNC-IVF or Sub-NC children, but this neurological disadvantage did not reach statistical significance (table II). In accordance, the prevalence of dysfunction in specific neurological domains was similar in the three groups. As the prevalence of complex MND was low (table II), we used in the multivariate statistical analysis MND, denoting the presence of either simple or complex MND, as outcome parameter. The multivariate analyses confirmed that neither the ovarian hyperstimulation (COH-IVF versus MNC-IVF), nor the *in vitro* laboratory procedures (MNC-IVF versus Sub-NC), nor the

Table III - Multivariate regression analyses of the influence of ART components and history of subfertility on neurological outcome

	Adjusted median difference (95% CI)	***P* - value**
Outcome measure:		
Fluency score - **log (14.5 - fluency score)**		
COH-IVF versus Sub-NC	0.006 (-0.027 – 0.039)[a]	0.700
MNC-IVF versus Sub-NC	0.007 (-0.026 – 0.040)[b]	0.676
COH-IVF versus MNC-IVF	0.003 (-0.033 – 0.038)[c]	0.888
Subfertile group versus reference group	0.039 (0.014 – 0.064)[d]	0.003
Outcome measure:		
Neurological optimality score - **ln (59.5 - NOS)**		
COH-IVF versus Sub-NC	-0.061 (-0.187 – 0.066)[a]	0.344
MNC-IVF versus Sub-NC	-0.011 (-0.137 – 0.115)[b]	0.867
COH-IVF versus MNC-IVF	-0.045 (-0.182 – 0.093)[c]	0.523
Subfertile group versus reference group	0.153 (0.062 – 0.245)[d]	0.001

	Adjusted odds ratio (95% CI)	***P* - value**
Outcome measure: MND		
COH-IVF versus Sub-NC	2.161 (0.575 – 8.122)[a]	0.254
MNC-IVF versus Sub-NC	1.221 (0.277 – 5.393)[b]	0.792
COH-IVF versus MNC-IVF	1.917 (0.518 – 7.100)[c]	0.330
Subfertile group versus reference group	0.386 (0.152 – 0.979)[d]	0.045

Note: Linear regression analyses were used to compare the fluency score and NOS between groups. Logistic regression analyses were used to compare the prevalence of MND between groups.

a Adjusted for gestational age, time to pregnancy, gestational age and vanishing twins. Bonferroni correction applied.

b Adjusted for gestational age and time to pregnancy. Bonferroni correction applied.

c Adjusted for gestational age, paternal age and vanishing twins. Bonferroni correction applied.

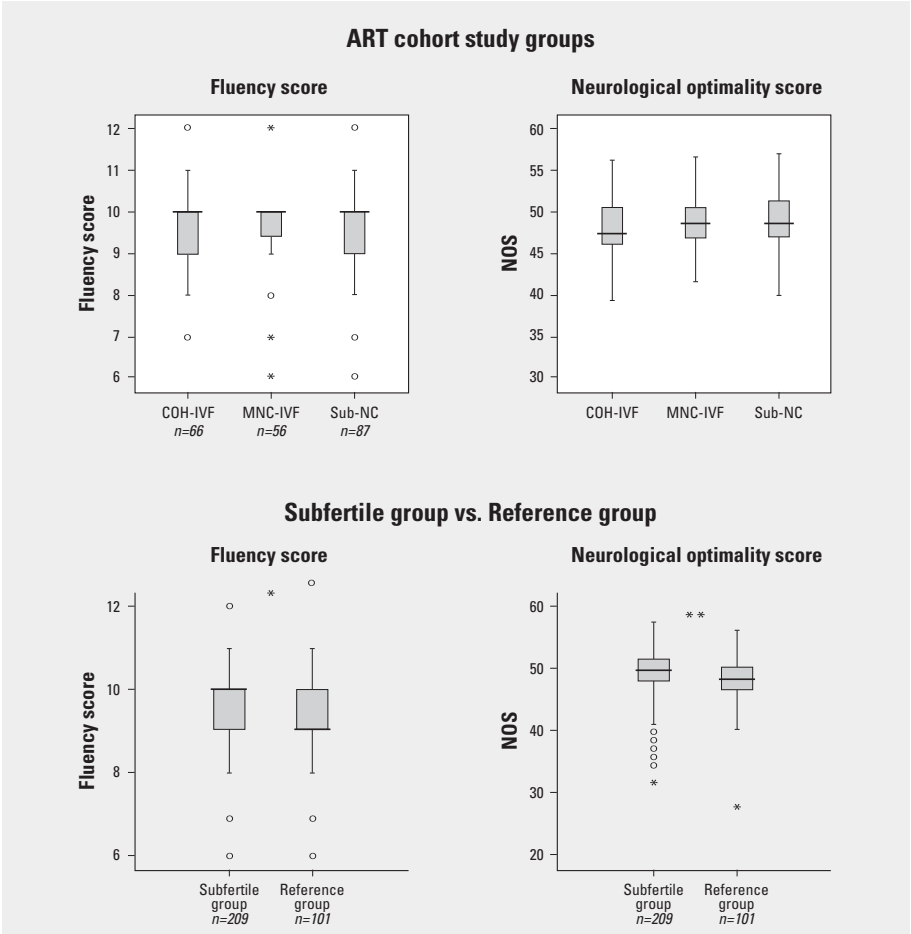
d Adjusted for gestational age, birthweight, firstborn, signs of fetal distress, caesarean section, breastfed for > 6 weeks, maternal age, education level mother, education level father and corrected age at examination at 2 years.

combination of both factors (COH-IVF versus Sub-NC) affected the occurrence of MND (table III).

The primary outcome parameter, the fluency score, and the other additional outcome parameter, the NOS, were similar in the three groups (figure 2). The similarity in fluency score and NOS in the three groups was confirmed in the multivariate analyses (table III).

Transforming the CI for the differences between the groups (table III) back into the original scale results in the following interpretation: assuming that the corrected median fluency score is 10 in the Sub-NC group, the CIs for corrected medians for group MNC-IVF and COH-IVF are both (9.7–10.4); assuming that the

Figure 2 - Fluency scores and NOS.



*Fluency score and NOS in the three ART cohort groups (upper panel) and the subfertile group and the reference group (lower panel). *P < .01, **P < .001.*

score is 10 for the MNC-IVF group, the CI for the COH-IVF group median is (9.6–10.4). For the NOS, assuming that the median score of 49 for the Sub-NC group results in CIs (47.5–50.1) and (46.8–49.7) for medians in the group MNC-IVF and COH-IVF, respectively; assuming that the score is 48 for the COH-IVF group, the CI for the MNC-IVF group is (45.7–46.9). The resemblance in neurological condition of the three groups allowed us to pool them to form the subfertile group.

Neurological condition of the subfertile and reference groups

Parents of 167 reference children were invited to participate in the study. Parents of 37 infants refused to participate. Logistic problems hampered the assessment of another 29 children. Eventually, 101 (61%) children were examined. Maternal

Table IV - Characteristics of parents and infants of the subfertile group and the reference group

Characteristics	Subfertile group[a]	Reference group	*P* - value
	n = 209	*n = 101*	
Male gender; n (%)	108 (52)	47 (47)	0.396
First born, n (%)	136 (65)	52 (52)	0.022
Birth characteristics:			
Gestational age (in weeks), median (range)	39.9 (30.1-42.7)	40.0 (35.7-42.4)	0.275
Birth weight in grams[f], mean (range)	3460 (575)	3615 (465)	0.022
Low birth weight (< 2500 gram), n (%)	12 (6)	0 (0)	0.021
Small for gestational age[b], n (%)	5 (2)	6 (6)	0.168
Preterm birth (< 37 weeks), n (%)	20 (10)	2 (2)	0.015
Signs of fetal distress[c], n (%)	73 (35)	20 (20)	<0.001
Caesarean section, n (%)	47 (23)	8 (8)	0.002
Neonatal characteristics:			
Breastfed for > 6 weeks[f], n (%)	98 (48)	59 (62)	0.030
Parental characteristics:			
Maternal age at conception, mean (range)	33.0 (22.2-40.9)	30.9 (18.0-40.4)	<0.001
Smoking during pregnancy (mother), n (%)	23 (11)	11 (11)	0.976
Parental socioeconomic status:			
Education level mother (high[d]), n (%)	83 (40)	53 (53)	0.034
Education level father (high[d])[f], n (%)	79 (39)	55 (56)	0.004
Time to pregnancy (categorical[e])[f]	4 (0-5)	0 (0-1)	<0.001
Corrected age at examination at 2 years of age (in months)	24.9 (23.2-30.2)	25.5 (23.5-29.1)	<0.001

Note: Mann-Whitney U tests or Student's t-test and χ^2 tests or Fisher's exact tests were used to compare between groups.

[a] Group COH-IVF, MNC-IVF and Sub-NC were pooled to form the subfertile group.

[b] Birth weight < 10%.

[c] Signs of fetal distress denoted by meconium stained amniotic fluid and/or cardiotocographic signs and/or acidosis.

[d] University education or vocational colleges.

[e] Time to pregnancy as a categorical value: 0 = 0 - ½ year, 1 = ½ - 1 year, 2 = 1 - 2 years, 3 = 2 -3 years, 4 = 3 - 5 years, 5 = > 5 years.

[f] Missing data in two groups: birth weight n=7, breastfed n=10, education level father n=4, time to pregnancy n=1. Note that in the percentages missing values have been taken into account.

education of non-participating children was significantly lower than that of participating children ($P < .01$), but gender distribution and gestational age at birth were similar for participants and non-participants. Table IV illustrates that perinatal and social characteristics of the subfertile and the reference groups differed considerably. Compared with the fertile reference group children of the subfertile group more often were firstborn ($P = .022$) and preterm ($P = .015$), had a lower birthweight ($P = .022$), showed more often signs of fetal distress ($P < .001$), more often were born following Caesarean section ($P = .002$) and were slightly younger at the follow-up assessment ($P < .001$). In addition, in the subfertile group maternal age was higher ($P < .001$) and parental educational level was lower (maternal: $P = .034$, paternal: $P = .004$) than in the reference group. Finally, parents of the subfertile group needed more time to get pregnant ($P < .001$).

Children of the subfertile group tended to show less MND than the reference children. In the univariate analysis the difference did not reach statistical significance (table II). Yet, when confounders were taken into account the difference was statistically significant ($P = .045$; table III). Furthermore, a minor difference was found in the specific domains of dysfunction: the reference group showed more often dysfunctional posture and muscle tone regulation than the subfertile group ($P < .01$; table II).

Similar results were found for the fluency score and the NOS. Both fluency score and NOS of the subfertile group were higher than those of the reference group [fluency score: median values 10 (6–12) and 9 (6–13), respectively; $P = .002$; NOS: median values 49 (32–57) and 47 (28–55); $P < .001$; figure 3]. Multivariate analyses confirmed that differences in fluency and NOS scores between the subfertile group and the reference group were statistically significant ($P = .003$ and $P = .001$, respectively; table III).

Transforming the CIs for the differences between both groups (table III) back into the original scale results in the following interpretation: assuming that the corrected median fluency score is 9 in the reference group, the CIs for corrected medians for the subfertile group is (9.2–9.8); assuming that the score is 10 for the subfertile group, the CI for the reference group median is (10.1–10.6). For the NOS, assuming that the median score of 47 for the reference group results in the CI (47.8–49.7) for the median in the subfertile group; assuming that the score is 49 for the subfertile group, the CI for the reference group is (49.6–51.3).

Underlying subfertility causes and neurological outcome

None of the specific causes of subfertility were related to the fluency score, NOS or the type or severity of MND (table II). This was confirmed in multivariate analyses (data not provided).

DISCUSSION

The present study indicates that neurological outcome in 2 year olds is not influenced by ovarian hyperstimulation, the in vitro laboratory procedure itself or a combination of both. The findings support the results of the Groningen ART cohort study at younger ages (Middelburg *et al.*, 2009a,b).

Our results are in line with most good quality studies in which children born after ART were followed until at least the age of 2 years (Brandes *et al.*, 1992; Agarwal *et al.*, 2005; Wikstrand *et al.*, 2006; Knoester *et al.*, 2007; Ludwig *et al.*, 2009a). They strengthen the notion that ART is not associated with adverse neurological outcome at early age since we used a standardised and sensitive neurological assessment which allows for the detection of subtle differences in outcome (Hadders-Algra, 2005). Our data also indicated that specific causes of subfertility, which may be associated with higher rates of spontaneous abortion (Sutcliffe and Ludwig, 2007), are not associated with worse neurological outcome.

Unexpectedly, we found that neurological outcome of the subfertile group was better than that of the reference group. Previously, we found that neurological condition of the subfertile group at 3 months was slightly worse than that of a reference group of the general population (Middelburg *et al.*, 2009a). This difference may be attributed to the selection of the reference groups. The reference group at 3 months consisted of a group which was representative for the general population—the children had been assessed as part of a general health check-up provided for all children. The current reference group consisted of 2 year olds whose parents volunteered to form a reference group for the ART cohort study. This may have introduced a selection bias. Also, Knoester *et al.* (2007) reported in the control group of naturally conceived children of fertile couples a prevalence rate of MND which was substantially higher than that in the general population (Hadders-Algra, 2010). Recently, Carson *et al.* (2010) elegantly demonstrated the importance of a proper comparison group: they showed that when the control group was not properly matched to the ART study group, children born after ART had a better cognitive performance than the reference group. But when the background factors of the reference group closely resembled that of the ART group differences in cognitive disabilities disappeared. Perinatal and social background factors of our reference group differed substantially from those of the Groningen ART cohort groups. From a neurodevelopmental view perinatal and social background factors in the reference group were more favourable than those in the subfertile group (table IV). Multivariate analyses were applied to adjust for the large differences in background between the two groups.

Strengths and limitations

One of the major strengths of the study is the composition of our study groups which allowed for a disentangling of the effect of two aspects of ART. Our study was able to demonstrate that neither ovarian hyperstimulation nor the in vitro procedure was associated with an increased risk for neurological non-optimality at the age of 2 years.

Parental characteristics of the naturally conceived control group, consisting of children born to subfertile couples (Sub-NC), closely resembled those of both COH-IVF and MNC-IVF groups. As a result, the effects of potential confounders and a potential overestimation of the effect of ART were minimized.

Another major strength of the present study is the application of a sensitive and age-specific technique to assess neurological condition. The strength of the Hempel assessment is illustrated by the study of Bouwstra *et al.* (2006), which demonstrated a negative effect of neonatal trans-fatty acid status on neurodevelopmental outcome using the assessment according to Hempel but not by using the Bayley's Scale of Infant Development. Although subtle differences in neurological outcome, such as a few points reduction in the NOS or fluency score, may not have clinical relevance for individual persons, minor deviations in neurodevelopmental outcome for substantial subpopulations may affect society at large, especially since the number of fertility problems is steadily rising.

The prospective design of our study, in which couples of the COH-IVF, MNC-IVF and Sub-NC groups were invited in the third trimester of pregnancy, reduced potential selection bias based on the child's health or development.

Additional strengths of our study are the minimal post-natal attrition (3%) and the 'blinding' of assessors to the mode of conception of the ART cohort. Whether blinding in ART studies is valuable, was questioned by Ludwig *et al.* (2009b) who found that the assessors' feeling about the mode of conception was correct in 75% of cases. The likelihood that the assessors in our study would guess the conception mode was minimized, because all parents of the ART cohort had experienced subfertility. However, we had been unable to prevent knowledge of conception mode of the assessor involved in the examination of the reference group.

It is a limitation of the study that we studied singletons only. This means that the results cannot be generalised to children born after multiple gestation. It is well known that controlled ovarian hyperstimulation IVF is associated with multiple birth and that being a member of a multiple is associated with an increased risk for developmental problems (Pinborg *et al.*, 2003).

A limitation of the study is the composition of the fertile reference group. Our results underscore the need of prospectively recruited control groups. Another

limitation of the study was the size of the MNC-IVF group, which was slightly smaller than the 64 needed for an adequate power of the study. However, the finding that outcome of the three subfertile groups was very similar, suggests that it is unlikely that larger groups would have revealed statistically significant differences between the groups. Owing to the relatively small number of children in our groups, the study was unable to evaluate the association between ART procedures and rare, severe neurodevelopmental disorders.

In conclusion, the neurological condition of 2 year olds born after COH-IVF or MNC-IVF is similar to that of peers born to subfertile couples who conceived without ART. This indicates that neither the ovarian hyperstimulation, nor the IVF laboratory procedures or a combination of both factors is associated with a worse neurological outcome up until the age of 2 years. Additionally, our data suggest that subfertility is not associated with a worse neurological outcome at 2 years. Although the findings of the present Groningen ART cohort study are reassuring, we have to keep in mind that subtle neurodevelopmental disorders may emerge when children grow older.

Part II
The Groningen
ART cohort study

6

Developmental status and behaviour at 2 years

M. Jongbloed-Pereboom
K.J. Middelburg
M.J. Heineman
A.F. Bos
M. Hadders-Algra

Fertility Sterility 2011; 95: 2283-2289.

ABSTRACT

Objective: To evaluate whether children's cognitive and psychomotor development and behaviour at 2 years is affected by ovarian hyperstimulation and the IVF laboratory procedures or subfertility.
Design: Prospective longitudinal cohort study
Setting: University Medical Center Groningen, the Netherlands
Patients: Singletons born after COH-IVF (n=66), MNC-IVF (n=56), singletons born to subfertile couples who conceived naturally (sub-NC, n=87) and a reference group of 101 2-year-old singletons born to fertile couples.
Intervention: None.
Main Outcome Measures: Bayley Scales of Infant Development and Achenbach's Child Behaviour Checklist (CBCL).
Results: Mental and psychomotor development and behavioural outcome in COH-IVF, MNC-IVF and Sub-NC groups was not different. Developmental outcome and behaviour of the subfertile groups was largely similar to that of the fertile reference group. Nevertheless, the subfertile groups scored higher on the scale of anxious-depressed behaviour than the reference group.
Conclusion: This present relatively small study found no differences in cognitive and psychomotor development and behaviour at 2 years in children born after COH-IVF, MNC-IVF or naturally conceived children of subfertile parents. Replication of the study is needed before firm conclusions can be drawn. Furthermore, long-term follow-up is needed to confirm these findings in older children.

INTRODUCTION

In 2006 up to 4% of children in European countries were born following ART (de Mouzon *et al.*, 2010). Because of the growing number of ART-conceived children, small changes in the children's development and behaviour are of importance to society. Up to now, results of most studies on the effect of ART on neurodevelopmental outcome have been reassuring (Sutcliffe and Ludwig, 2007; Middelburg *et al.*, 2008), nevertheless many studies are hampered by methodological shortcomings (Middelburg *et al.*, 2008). Concern on developmental outcome is justified as it is known that ART is associated with a higher risk for adverse perinatal outcomes, including preterm birth and low birth weight (Helmerhorst *et al.*, 2004; Jackson *et al.*, 2004). Children born preterm or with low birth weight have an increased risk for developmental problems (Bhutta *et al.*, 2002; Moster *et al.*, 2008).

Different components of the ART-procedure may change embryo development and may influence the development of the conceived child. Suggested points of concern are, for instance, ovarian hyperstimulation, the effects of laboratory procedures involved in the in vitro fertilisation (IVF) or intracytoplasmic sperm injection (ICSI), and consequences of vanishing twins (Sutcliffe and Ludwig, 2007; Jackson *et al.*, 2004; Olivennes *et al.*, 1993; Pinborg *et al.*, 2005; Kapiteijn *et al.*, 2006; Griesinger *et al.*, 2008). Furthermore, parental characteristics associated with subfertility may affect child development (Sutcliffe and Ludwig, 2007; Jackson *et al.*, 2004; Olivennes *et al.*, 1993; Draper *et al.*, 1999).

The Groningen ART-cohort was initiated to study the potential effect of different components of the ART procedure. Three groups of children were recruited prospectively to disentangle the effects of ovarian hyperstimulation and the in vitro procedure itself. The first group consists of children born following a conventional, so called 'controlled-ovarian hyperstimulation'-IVF procedure (COH-IVF). The second group are children born following IVF in the modified natural cycle (MNC-IVF). In MNC-IVF medication use is minimal and the aim is to use the one single oocyte that develops to dominance naturally (Nargund *et al.*, 2007). The third group consists of naturally conceived children born to subfertile couples (sub-NC). Potential differences in outcome of COH-IVF and MNC-IVF children may be attributed to ovarian hyperstimulation and the comparison of MNC-IVF children and sub-NC children was used to study the effect of the in vitro procedure. In the current study we additionally recruited a reference group of 2-year-old children born to fertile parents to study the effect of subfertility.

Previously our group reported on the neurodevelopmental outcome of singletons at the ages of two weeks, three, four, ten and eighteen months (Middelburg *et al.*, 2009; Middelburg *et al.*, 2010). At those ages neurodevelopment

of COH-IVF, MNC-IVF and sub-NC children was similar. However, at the age of 3 months the three subfertile groups showed a less optimal neurodevelopment than a reference group representative for the general population (Middelburg *et al.*, 2010). Since subtle neurodevelopmental disorders may emerge when children grow older (Hadders-Algra, 2002), continuation of follow-up is needed.

In the present study, we report cognitive and motor development and child behaviour of singletons in the Groningen-ART cohort and in a reference population at the age of 2 years. We address the question whether these parameters of neurodevelopmental outcome are related to ovarian hyperstimulation, the in vitro procedure itself, a combination of these two factors or a history of subfertility.

METHODOLOGY

Participants

From March 2005 to December 2006 we invited all couples who achieved a singleton pregnancy after IVF/ICSI at the University Medical Center Groningen (UMCG) to participate in a longitudinal study on neurodevelopmental outcome of IVF/ICSI children (Middelburg *et al.*, 2010). This resulted in two groups. The first group consisted of singletons born after IVF/ICSI in the controlled ovarian hyperstimulation cycle (COH-IVF, n=68) and the second group of singletons born after IVF/ICSI in the modified natural cycle (MNC-IVF, n=57). Inclusion criteria for MNC-IVF treatment are reported by Pelinck *et al.* (2008). Excluded were children born after cryopreservation or donation of oocytes or embryos. A third group was formed by singletons born to couples who achieved pregnancy naturally while waiting for fertility evaluation or treatment (subfertile naturally conceived, sub-NC, n=90). These couples had tried to achieve pregnancy for at least one year.

A fourth, additional reference group was recruited from February until October 2009 at five child welfare clinics in the northern part of the Netherlands. All parents of singleton two-year-old children visiting the child welfare clinic for routine general health care were invited to participate. Exclusion criteria were any form of assisted conception and a time to pregnancy of more than one year.

The study was approved by the ethics committee of the UMCG. Parents provided written informed consent for participation of their children in the study.

Procedure

For children in the Groningen-ART cohort detailed information on fertility, obstetric and social characteristics was available. Similar information of the reference

population was collected by means of a questionnaire at the time of visit. The neurodevelopmental assessment was carried out by the first author and four trained research assistants. They were blind to mode of conception in the Groningen ART cohort study, but were aware of the fact whether a child was recruited at a child welfare centre or not.

Measures

Mental and Psychomotor development

Mental and motor development was tested with the Dutch version of the Bayley Scales of Infant Development, second edition (BSID-II-NL (Van der Meulen *et al.*, 2004)). The mental developmental index (MDI) is determined by performance on items that measure visual and auditory information processing, eye-hand coordination, imitation, language skills, memory and problem solving. The psychomotor developmental index (PDI) consists of items assessing gross and fine motor function. Both the MDI and PDI have a mean value of 100 and a standard deviation of 15.0 (Van der Meulen *et al.*, 2004). Reliability and validity of the BSID-II-NL are satisfactory (Ruiter *et al.*,2008).

Child behaviour

Child behaviour was measured with the Achenbach Child Behaviour Checklist (CBCL) for children aged 1.5 to 5 years (Achenbach and Rescorla, 2000). The CBCL consists of 100 items concerning emotional and behavioural problems, which are scored by parents, based on the preceding 2 months. Raw scores are normalised into T-scores with a mean value of 50 and a standard deviation of 10. Good reliability and validity have been reported for the CBCL (Achenbach and Rescorla, 2000).

Statistical analyses

Power calculation of the study was based on neurological outcome at the age of 18 months (Middelburg *et al.*, 2009). For the current study, to detect a difference of at least half a standard deviation on the Bayley MDI or PDI (mean 100, standard deviation 15), with 80 % power, at least 64 children had to be included per group.

Data analysis was performed with the Statistical Package for Social Sciences (SPSS) 15.0 for Windows. Group differences were evaluated by means of Student's t-tests, Mann Whitney tests, and χ^2 tests. Multivariate linear regression analyses were performed to further test for group differences while taking into account confounders. Variables considered as confounders were variables for which groups differed at 5% significance level in the univariate analyses. In addition, gestational age, gender and mother's educational level was entered in the multivariate analyses, since we know from literature that these are important

predictors for developmental outcome. Bonferroni correction was applied in the analysis of the outcome measures between the COH-IVF group, MNC-IVF group and sub-NC group, but not in the analysis of the background factors or in the analyses between the subfertile cohort and the reference group.

RESULTS

Two hundred and nine children (97%) of 215 eligible ART-cohort children participated in the follow-up study at two years of age. The parents of four children lost interest in the study, another couple could not be traced after change of address and one child (MNC-IVF group) died in the neonatal period. The BSID-II-NL could not be administered in one sub-NC child due to severe developmental delay. Parents of six children did not return the CBCL.

Parents of 167 reference children were invited to participate in the study. Parents of 61 infants refused to participate and logistic problems hampered the assessment of another 12 children. Eventually, 101 (60 %) children were examined. Maternal education of non-participating children was significantly lower than that of participating children ($P < .01$), but gender distribution and gestational age at birth were similar for participants and non-participants. One reference child was so shy that it precluded the administration of the BSID-II-NL. Parents of two reference children did not return the CBCL.

Background characteristics of the children in the COH-IVF-, MNC-IVF group and the sub-NC-group are shown in table I. Overall, characteristics of the three subfertile groups were similar. However, birthweight of the COH-IVF singletons was significantly lower than birthweight of the sub-NC children, and paternal age was significantly higher in the COH-IVF group than in the MNC-IVF group. Furthermore, time to pregnancy was significantly longer in the two IVF groups than in the sub-NC group and - obviously - vanishing twins occurred more often in the COH-IVF group than in the MNC-IVF group and the sub-NC group.

Cognitive, motor and behavioural outcome in the COH-IVF, MNC-IVF and sub-NC group is presented in table II. The various outcome parameters of neurodevelopmental outcome were similar in the three groups, except for the PDI, which was significantly lower in the COH-IVF group than in the MNC-IVF group after correction for multiple testing. In the multivariate analyses (table III) this difference was no longer significant. The similarity in outcome allowed us to pool the data of the children of the COH-IVF, MNC-IVF and sub-NC groups to form a subfertile group. Results of the subfertile group were compared with the fertile reference group.

Table 1 - Infant and parental characteristics of COH-IVF, MNC-IVF and sub-NC children, the subfertile cohort and the reference group.

Characteristics	COH-IVF *(n = 209)*[a]	MNC-IVF *(n = 66)*	subNC *(n = 56)*	Subfertile cohort *(n = 87)*	Reference group *(n = 101)*
Sex: male, no. (%)	36 (55)	27 (48)	45 (52)	108 (52)	47 (47)
Firstborn, no. (%)	45 (68)	38 (68)	53 (61)	136 (65)b	51 (50)[b]
Corrected age (mo) at examination, median (range)	24.9 (23.3-30.2)	24.9 (23.3-27.9)	25.0 (23.2-28.9)	24.9 (23.2-30.2)[c]	25.5 (23.5-29.1)[c]
Birth characteristics					
Preterm birth (<37 wk), no. (%)	7 (11)	6 (11)	7 (8)	20 (10)[b]	2 (2)[b]
Birth weight (g), median (range)	3395 (1980-4700)[d]	3405 (2170-4680)	3620 (1150-4710)[d]	3460 (1150-4710)[b]	3660 (2700-4600)[b,e]
Neonatal characteristics					
Apgar score (5min) <7; no. (%)	0 (0)	0 (0)[e]	1 (1)[e]	--	--
Neonatal intensive care admission, no. (%)	1 (2)	2 (4)	6 (7)	9 (4)	1 (1)
Parental characteristics					
Maternal age (y), median (range)	32.9 (26.3-40.9)	32.8 (25.3-37.5)	33.3 (22.2-40.3)	33.0 (22.2-40.9)[c]	30.9 (18.0-40.4)[c]
High maternal education level, no. (%)[f]	22 (33)	21 (38)	40 (46)	83 (40)	53 (52)
Ethnicity mother (white), no. (%)	62 (94)[e]	55 (98)	82 (94)	199 (96)	95 (94)
Smoking during pregnancy, no. (%)	7 (11)	7 (13)	9 (10)	23 (11)	11 (11)
Paternal age (y), median (range)	35.7 (27.5-56.1)[e,g]	34.3 (28.3-47.8)[e,g]	35.4 (25.2-52.6)	--	--
High paternal education level, no. (%)[f]	29 (46)[e]	18 (32)[e]	32 (37)	79 (39)[h]	56 (57)[e,h]
Ethnicity father (white), no. (%)	60 (91)	55 (98)	78 (90)	193 (95)	90(89)
Fertility parameters:					
ICSI, no. (%)	42 (64)	28 (50)	na	--	--
Time to pregnancy (y), median (range)[i]	4.0 (0.1-13.3)[j]	3.8 (0.1-13.2)[k]	2.1 (0.1-11.3)[j,k]	--	--
Vanishing twins, no. (%)	8 (12)[g,j]	1 (2)[g]	0 (0)[j]	--	--

Note: NA = not applicable.

[a] The subfertile cohort is the total of COH-IVF, MNC-IVF, and subNC groups.

[b] Significantly different, subfertile cohort vs. reference group, P <.05.

[c] Significantly different, subfertile cohort vs. reference group, P <.001.

[d] Significantly different, COH-IVF vs. subNC, P <.05.

[e] Missing values: for the COH-IVF group there was one missing value of ethnicity of mother, two for paternal age, and three for paternal education level. For the MNC-IVF group there was one missing Apgar score, two missing values for paternal age, and three for paternal education level. For the subNC group there were three missing Apgar scores and one missing value of time to pregnancy. For the reference population there were seven and three missing values for birth weight and education level of father, respectively.

[f] University education or vocational colleges.

[g] Significantly different, COH-IVF vs. MNC-IVF, P <.05.

[h] Significantly different, subfertile cohort vs. reference group, P <.01.

[i] Time to pregnancy was defined as the period between the onset of attempts to conceive or a previous pregnancy and the last menstrual period before pregnancy.

[j] Significantly different, COH-IVF vs. subNC, P <.001.

[k] Significantly different, MNC-IVF vs. subNC, P <.01.

Background characteristics of the subfertile group and the fertile reference group are shown in table I. In the subfertile group the number of first born and preterm children was higher, birth weight was lower, maternal age was higher, and

parental education and the child's age at assessment were lower than in the fertile reference group.

The results on the cognitive, motor and behavioural tests of the subfertile group and the fertile reference group are presented in table II. The univariate analyses did not reveal differences between the groups. But, in the multivariate analyses (table III) a difference emerged: children of the subfertile group scored higher on the CBCL's anxious-depressed scale than their peers of the reference group. Nevertheless all values were well within normal range.

We performed additional analyses without taking into account gestational age and birthweight as they may be considered as mediators on the pathway from assisted conception to developmental outcome. The re-analyses indicated no differences between the three groups, or between the subfertile and the reference group, including no difference in anxious-depressed behaviour.

Table II - BSID-II and CBCL scores in children born following COH-IVF, MNC-IVF and sub-NC children, the subfertile cohort and the reference group.

	COH-IVF	MNC-IVF	subNC	Subfertile cohort	Reference group
BSID-II-NL (n)	66	56	86	208	100
MDI, mean (SD)	98.0 (13.1)	101.0 (12.0)	99.8 (12.9)	99.6 (12.8)	100.8 (15.0)
PDI, mean (SD)	86.1 (13.4)	93.0 (15.3)	89.5 (15.9)	89.4 (15.2)	92.9 (14.7)[a]
CBCL (n)	66	55	82	203	99
Total problems scale, mean (SD)	46.1 (8.1)	46.9 (8.8)	48.6 (8.7)	47.3 (8.6)	46.6 (8.5)
Internalizing problems, mean (SD)	43.4 (9.0)	44.4 (8.9)	46.2 (8.2)	44.8 (8.7)	44.9 (8.8)
Externalizing problems, mean (SD)	49.7 (7.6)	50.3 (9.6)	51.7 (9.5)	50.7 (9.0)	49.1 (8.8)
Emotionally reactive, mean (SD)	52.6 (4.6)	52.9 (4.2)	53.8 (5.3)	53.2 (4.8)	52.8 (4.8)
Anxious/depressed, mean (SD)	50.7 (2.4)	50.5 (1.0)	50.7 (1.6)	50.6 (1.8)	51.0 (2.5)
Somatic complaints, mean (SD)	53.3 (5.8)	53.7 (5.1)	53.3 (5.6)	53.4 (5.5)	53.7 (5.3)
Withdrawn, mean (SD)	51.9 (4.3)	51.6 (3.8)	52.5 (4.2)	52.1 (4.2)	52.0 (3.4)
Sleep problems, mean (SD)	52.0 (4.3)	51.7 (2.9)	53.9 (6.9)	52.7 (5.4)	53.2 (6.2)
Attention problems, mean (SD)	54.5 (5.3)	55.5 (6.8)	54.9 (5.6)	54.9 (5.9)	53.5 (5.5)
Aggressive behaviour, mean (SD)	53.1 (4.0)	53.8 (4.0)	54.7 (6.2)	53.9 (5.1)	53.1 (5.5)

Note: Internalizing problems combines the syndrome scales emotionally reactive, anxious/depressed, somatic complaints, and withdrawn. Externalizing problems consists of the syndrome scales attention problems and aggressive behavior. A total problems score was calculated by summing scores of all items.

[a] One missing value Bayley PDI reference group.

Table III - Multivariate regression analyses of contribution of IVF method or subfertility on outcome measures.

Compared Groups	Mean Difference (95% CI)	*P* value	Adjusted mean difference (95% CI)	*P* value
MDI				
COH-IVF versus MNC-IVF	-3.06 (-7.60 - 1.48)	.18	-1.86 (-6.65 – 2.93)	.44
COH-IVF versus SubNC	-1.86 (-6.07 – 2.35)	.38	-1.73 (-6.25 – 2.79)	.45
MNC-IVF versus SubNC	1.20 (-3.07 – 5.48)	.58	1.47 (-2.78 – 5.72)	.50
Subfertile versus reference	-1.28 (-4.52 – 1.96)	.44	-0.51 (-4.12 – 3.09)	.78
PDI				
COH-IVF versus MNC-IVF	-6.94 (-12.09 - -1.79)[a]	.01	-5.46 (-10.95 – 0.03)	.05
COH-IVF versus subNC	-3.47 (-8.28 – 1.34)	.16	-2.63 (-7.77 – 2.52)	.31
MNC-IVF versus subNC	3.47 (-1.85 – 8.79)	.20	3.57 (-1.72 – 8.85)	.18
Subfertile versus reference	-3.52 (-7.13 – 0.09)	.06	1.47 (-2.66 – 5.60)	.49
CBCL total problems				
COH-IVF versus MNC-IVF	-0.78 (-3.83 – 2.27)	.61	-1.07 (-4.36 – 2.23)	.52
COH-IVF versus subNC	-2.54 (-5.31 – 0.22)	.07	-3.01 (-6.11 – 0.09)	.06
MNC-IVF versus subNC	-1.76 (-4.78 – 1.26)	.25	-1.58 (-4.75 – 1.58)	.32
Subfertile versus reference	0.79 (-1.28 – 2.85)	.46	-0.63 (-3.07 – 1.81)	.61
CBCL internalizing problems				
COH-IVF versus MNC-IVF	-1.01 (-4.24 – 2.23)	.54	-1.64 (-5.16 – 1.89)	.36
COH-IVF versus subNC	-2.83 (-5.63 - -0.03)	.05	-3.17 (-6.26 - -0.07)	.05
MNC-IVF versus subNC	-1.82 (-4.75 – 1.11)	.22	-1.42 (-4.48 – 1.65)	.36
Subfertile versus reference	-0.13 (-2.24 – 1.98)	.90	0.78 (-1.73 – 3.29)	.54
CBCL externalizing problems				
COH-IVF versus MNC-IVF	-0.51 (-3.63 – 2.61)	.75	-0.42 (-3.72 – 2.87)	.80
COH-IVF versus subNC	-1.95 (-4.82 – 0.91)	.18	-2.65 (-5.86 – 0.57)	.11
MNC-IVF versus subNC	-1.44 (-4.74 – 1.11)	.22	-1.32 (-4.79 – 2.15)	.45
Subfertile versus reference	1.54 (-0.62 – 3.70)	.16	1.70 (-4.22 – 0.82)	.18
CBCL emotionally reactive				
COH-IVF versus MNC-IVF	-0.25 (-1.86 – 1.36)	.76	-0.49 (-2.26 – 1.29)	.59
COH-IVF versus subNC	-1.14 (-2.78 – 0.51)	.18	-0.97 (-2.81 – 0.88)	.30
MNC-IVF versus subNC	-0.88 (-2.58 – 0.82)	.31	-0.54 (-2.30 – 1.22)	.54
Subfertile versus reference	0.32 (-0.84 – 1.48)	.59	-0.26 (-1.66 – 1.14)	.71
CBCL anxious/depressed				
COH-IVF versus MNC-IVF	0.18 (-0.49 – 0.85)	.60	-0.04 (-0.74 – 0.67)	.92
COH-IVF versus subNC	-0.06 (-0.70 – 0.59)	.87	-0.07 (-0.79 – 0.65)	.85
MNC-IVF versus subNC	-0.24 (-0.72 – 0.25)	.34	-0.13 (-0.63 – 0.37)	.60
Subfertile versus reference	-0.35 (-0.85 – 0.14)	.16	0.67 (0.08 – 1.26)[a]	.03
CBCL somatic complaints				
COH-IVF versus MNC-IVF	-0.42 (-2.41 – 1.56)	.68	-0.67 (-2.86 – 1.52)	.55
COH-IVF versus subNC	-0.01 (-1.86 – 1.85)	.99	-0.06 (-2.12 – 2.00)	.95
MNC-IVF versus subNC	0.42 (-1.45 – 2.29)	.66	0.50 (-1.47 – 2.46)	.62
Subfertile versus reference	-0.31 (-1.63 – 1.00)	.64	0.39 (-1.16 – 1.93)	.62

▶▶▶

Table III - *Continued*

Compared Groups	Mean Difference (95% CI)	*P* value	Adjusted mean difference (95% CI)	*P* value
CBCL withdrawn				
COH-IVF versus MNC-IVF	0.32 (-1.17 – 1.81)	.67	-0.37 (-2.93 – 1.33)	.64
COH-IVF versus subNC	-0.56 (-1.96 – 0.84)	.43	-0.96 (-2.53 – 0.61)	.23
MNC-IVF versus subNC	-0.88 (-2.29 – 0.53)	.22	-0.90 (-2.38 – 0.57)	.23
Subfertile versus reference	0.07 (-0.88 – 1.02)	.89	0.23 (-0.90 – 1.36)	.69
CBCL sleep problems				
COH-IVF versus MNC-IVF	0.31 (-1.05 – 1.67)	.65	0.21 (-1.29 – 1.71)	.79
COH-IVF versus subNC	-1.88 (-3.81 – 0.06)	.06	-2.18 (-4.34 - -0.03)	.05
MNC-IVF versus subNC	-2.19 (-4.42 – -0.23)	.03	-2.24 (-4.26 - -0.22)	.03
Subfertile versus reference	-0.52 (-1.88 – 0.85)	.46	0.47 (-1.16 – 2.10)	.57
CBCL attention problems				
COH-IVF versus MNC-IVF	-1.07 (-3.26 – 1.12)	.33	-0.64 (-3.01 – 1.74)	.60
COH-IVF versus subNC	-0.44 (-2.24 – 1.37)	.63	-0.80 (-2.83 – 1.23)	.44
MNC-IVF versus subNC	0.64 (-1.48 – 2.76)	.55	0.74 (-1.47 – 2.94)	.51
Subfertile versus reference	1.39 (-0.01 – 2.78)	.05	-1.07 (-2.70 – 0.55)	.20
CBCL aggressive problems				
COH-IVF versus MNC-IVF	-0.71 (-2.16 – 0.75)	.34	-0.86 (-2.44 – 0.72)	.29
COH-IVF versus subNC	-1.67 (-3.43 – 0.10)	.06	-1.73 (-3.69 – 0.23)	.08
MNC-IVF versus subNC	-0.96 (-2.85 – 0.93)	.32	-0.86 (-2.82 – 1.11)	.39
Subfertile versus reference	0.83 (-0.43 – 2.09)	.20	-1.00 (-2.48 – 0.48)	.19

Note: COH-IVF vs. MNC-IVF adjusted for maternal education, sex, gestational age, vanishing twins, and paternal age. COH-IVF vs. SubNC adjusted for maternal education, sex, gestational age, vanishing twins, birth weight, and time to pregnancy. MNC-IVF vs. SubNC adjusted for maternal education, sex, gestational age, and time to pregnancy. Subfertile cohort vs. reference group adjusted for maternal education, sex, gestational age, birth weight, firstborn, age at examination, maternal age, and paternal education. Excluding maternal age from this analysis did not change the results. Substitution of gestational age by being born preterm or term did not change the results of the regression analyses.

CI = confidence interval.

[a] Significantly different, P <.05 for subfertile cohort vs. reference group, P <.017 comparing subfertile groups.

DISCUSSION

The present study shows no difference in development and behaviour of singletons born after COH-IVF, MNC-IVF and naturally conceived singletons born to subfertile couples. Nor did developmental outcome and behaviour of children born to subfertile couples differ from that of a reference group of children born to fertile couples.

One of the strengths of our study is the inclusion of the two IVF groups, the sub-NC group and the reference group. Inclusion of the two IVF groups allowed us to evaluate associations between ovarian hyperstimulation and child development and behaviour, while the sub-NC group allowed for the assessment of the effect of

the IVF-laboratory procedures with or without ovarian hyperstimulation. Finally, the reference population allowed us to study – to some extent - the effect of subfertility.

A second strength is the high participation and low attrition rate in the subfertile group. We reported previously that about 70 % of eligible couples decided to participate in the follow-up study during the third trimester of pregnancy. Participation was non-selective for relevant social and biological risk factors (Middelburg *et al.*, 2010). Further attrition after child birth was limited to 3 percent.

The study has several limitations. First of all, this study was slightly underpowered and due to the relatively small sample size it was not able to identify potential differences in outliers, i.e. in infants with serious neuropsychiatric problems. For this reason, it is not possible to make definite conclusions about differences between the groups. Second, the CBCL was filled in by parents. Information provided by parents may be vulnerable to social desirability, as being grateful for having a child may cause overwhelming happiness and may hinder the expression of feelings of anxiety and stress (Fisher *et al.*, 2008; Repokari *et al.*, 2005). Additionally, women who had a child after IVF might express feelings and fears less than women who conceived naturally (Ulrich *et al.*, 2004). The assessors were blind to mode of conception of the children of the ART cohort, but were aware of the fact whether a child was recruited at a child welfare centre or not. The potential assumption bias may be considered as a limitation of the study. Furthermore, we observed selective participation in the fertile reference group; children of mothers with a higher education were more likely to participate. This means that the fertile group is not representative for the general population, a problem which can not be prevented in studies which rely on parental cooperation. In the multivariate analyses we adjusted for most evident confounders, such as maternal education. However, multivariate analyses cannot overcome the problem of selection bias (Carson *et al.*, 2010). Finally, the selection criteria of the MNC-group may be considered as another limitation of the study. Selection could have resulted in better outcome in the MNC-group than in the COH-group. However, outcome in both groups was similar, also when confounders (selection criteria) were taken into account.

We studied singletons only. Therefore, results can not be generalised to children born after multiple gestation. It is well known that controlled ovarian hyperstimulation IVF is associated with multiple birth and that being a member of a multiple is associated with an increased risk for developmental problems (Pinborg *et al.*, 2003).

The finding that motor and mental development at 2 years was not different in children born after COH-IVF, MNC-IVF and sub-NC children is in line with our findings at earlier ages (Middelburg *et al.*, 2009; Middelburg *et al.*, 2010). However,

we previously found that neurological outcome of the three subfertile groups at 3 months was less optimal than that of a reference population representative of the general population (Middelburg *et al.*, 2010). At 2 years we did not find developmental differences between the subfertile cohort and the fertile reference group, except for a minor difference on the CBCL's anxious-depressed scale, which disappeared when gestational age and birth weight were not taken into account in the multivariate analysis.

Interestingly, we found in all four groups a discrepancy between performance on the MDI and PDI: PDI-scores were about ten points lower. Other studies have reported similar discrepancies (Gibson *et al.*, 1998; Bowen *et al.*, 1998). It has been suggested that the prevalence of non-optimal neuromotor performance and delayed motor development has increased recently (Hadders-Algra, 2007; Fleuren *et al.*, 2007). Lower PDI scores could be associated with higher levels of maternal warmth and sensitivity, whereas negative affection could represent a parenting style that promotes activity and motor behaviour (Pridham *et al.*, 2002; Treyvaud *et al.*, 2009). We do not know whether parenting style contributed to the PDI of the participants in our study.

Conclusion

In the present relatively small study we found no differences in cognitive and psychomotor development and behaviour in 2 year old singletons born after COH-IVF, MNC-IVF and naturally conceived singletons born to subfertile couples. Replication of the study is needed before firm conclusions can be drawn. In addition, it is important to note that developmental disorders may emerge when children grow older, as dysfunctions in more complex neurobehavioural functions first get expressed at school age (Hadders-Algra, 2002). Therefore, long-term follow- up of children born after IVF is required to confirm these findings

Part III - Follow-up of children born after IVF with Preimplantation Genetic Screening

7

Neurological condition of infants born after IVF with preimplantation genetic screening

K.J. Middelburg
M.J. Heineman
M.L. Haadsma
A.F. Bos
J.H. Kok
M. Hadders-Algra

Pediatric Research 2010; 67: 430-434.

ABSTRACT

Objective: Aim of this study was to evaluate the effect of preimplantation genetic screening (PGS) on neurodevelopmental outcome in children.
Methods: We conducted a prospective follow-up study of children born to women randomly assigned to in vitro fertilisation with or without PGS. Primary outcome was adverse neurological outcome at 18 months; secondary outcomes were types of minor neurological dysfunction (MND), neurological outcome before 18 months, neonatal intensive care admission, and congenital malformations.
Results: Twenty women in the PGS group participated with 25 children and 26 women in the control group participated with 31 children. Five PGS pregnancies (25%) and four control pregnancies (15%) resulted in birth of at least one child with an adverse neurological outcome (adjusted odds ratio: 2.3 [0.4–12.0]). Dysfunction in fine motor abilities and posture and muscle tone dysregulation tended to be present more frequently after PGS. Neurological outcome before 18 months, neonatal intensive care admission, and prevalence of congenital malformations were similar in study and control pregnancies. Nevertheless, at child level, rates of adverse outcome were higher after PGS.
Discussion: In conclusion, outcome in pregnancies after in vitro fertilisation (IVF) with and without PGS was similar. The small sample size precludes the conclusion that PGS is not associated with less favourable neurological outcome. Safety of new assisted reproductive techniques should be evaluated before large-scale implementation.

INTRODUCTION

Children born after assisted reproduction represent a sizeable part of the population; therefore, their health is of general concern. Nowadays, in Europe and the United States, 1–4% of children are born after assisted reproduction, and the numbers are still increasing (Andersen *et al.*, 2008;Wright *et al.*, 2008).

One of the major goals for new reproductive techniques is to enhance efficiency of assisted reproduction, and newer methods are regularly introduced to achieve this goal. One of these methods is in vitro fertilisation (IVF) with preimplantation genetic screening (PGS) for aneuploidy. In this procedure, embryos obtained with IVF are biopsied, which implies that a hole is made in the zona pellucida with laser or by chemical means. One or two blastomeres are aspirated, so that copy numbers of several sets of chromosomes can be determined. The concept behind this procedure is to identify and discard embryos with an abnormal chromosomal constitution because these might have a lower implantation potential or eventually lead to miscarriage (Wilton, 2002). Higher ongoing pregnancy rates were expected due to selection of the most viable embryos but have not been demonstrated in recent randomised controlled trials (RCTs) (Staessen *et al.*, 2004;Jansen *et al.*, 2008;Schoolcraft *et al.*, 2008;Staessen *et al.*, 2008). In fact, two trials have shown a reduction in ongoing pregnancies (Mastenbroek *et al.*, 2007;Hardarson *et al.*, 2008).

Besides efficacy, another important issue in assisted reproduction is the safety of the procedure for the developing foetus. Despite the invasiveness of PGS, information on developmental outcome after this procedure is scarce (Nekkebroeck *et al.*, 2008a;Nekkebroeck *et al.*, 2008b;Banerjee *et al.*, 2008), and information on neurological outcome is completely lacking. Therefore, we conducted a prospective, assessor-blinded follow-up study of children born to women randomly assigned to IVF with or without PGS. The aim of our study was to investigate the effect of PGS on the child's neurodevelopmental outcome. The primary outcome measure of this study was adverse neurological outcome at the age of 18 months. Secondary outcome measures were specific types of minor neurological dysfunction (MND) at 18 months, neurological outcome before 18 months, admission to a neonatal intensive care unit, and congenital malformations.

METHODOLOGY

Eligible for this follow-up study were children of women participating in one centre (University Medical Center Groningen) of a multicenter trial on the efficiency of PGS to establish ongoing pregnancies (Mastenbroek *et al.*, 2007). In the trial, exclusion criteria for women were age of IVF candidate younger than 35 or older than 41 year, previously failed IVF cycles, and objections against a possible double-embryo transfer. Randomisation of women was performed centrally with minimization for age (35–37 and 38–41 year) and reproductive technique (IVF and intracytoplasmic sperm injection), with stratification according to study centre before the start of the IVF procedures. Information concerning IVF treatment procedures has been reported previously (Mastenbroek *et al.*, 2007). For this study, women who conceived naturally during the trial were excluded from the follow-up study because the aim of our study was to investigate the effect of PGS on child outcome in IVF treatment. The protocol of the follow-up study was approved by the Medical Ethics Committee of the University Medical Center Groningen.

Couples were invited for the follow-up study during the third trimester of pregnancy. After written informed consent, inclusion took place during the first 2 weeks after birth. At the first appointment, demographic information (including e.g. parity, gestational age, birth weight, neonatal intensive care unit admission, parental age, and educational level) was collected on standardised charts. Time to pregnancy was retrieved from fertility charts. Information on child health up to 18 months was collected by history taking during assessments. Major congenital malformations were classified as malformations that generally cause functional impairment or require surgical correction (Bonduelle *et al.*, 2002). All other malformations were classified as minor.

Neurodevelopmental assessments

Follow-up consisted of standardised, age-specific neurological assessments at the ages of 2 weeks and 3, 4, 10, and 18 months post term. Age-specific testing is necessary in children due to abundant structural and functional changes in the nervous system, which induce changes in expression and prevalence of neurological dysfunction (Hadders-Algra, 2005). At 2 weeks and 3 months, quality of general movements (GMs) was assessed. Four classes of GM quality can be distinguished, being normal-optimal, normal-suboptimal, mildly abnormal, and definitely abnormal GMs (Hadders-Algra *et al.*, 2004). At 4 and 10 months, the Touwen Infant Neurological Examination was used, which resulted in the classification of normal, normal-suboptimal, MND, and definitely abnormal (Touwen, 1976).

The neurological examination according to Hempel was used at 18 months

(Hempel, 1993). In this assessment, children perform various motor tasks while playing in a standardised free-field situation. Functionality is tested in five different domains: fine motor function, gross motor function, posture and muscle tone, reflexes, and visuomotor function. Multiple signs in a domain result in a dysfunctional cluster. Children are classified as neurologically normal, simple MND, complex MND, or major neurological dysfunction based on the number of dysfunctional clusters (Hadders-Algra, 2003). Neurologically normal implies the presence of no dysfunctional clusters or only the presence of the cluster reflexes. Simple MND means the presence of one cluster of dysfunction, i.e. the isolated presence of fine motor, gross motor, visuomotor dysfunction, or mild dysregulation of posture and muscle tone. It is considered to reflect a normal, but non optimal form of brain function. Complex MND denotes the presence of two or more dysfunctional clusters; it is the form of MND with clinical relevance due to its clear association with learning and behavioural disorders (Hadders-Algra, 2002;Batstra *et al.*, 2003). Major neurological dysfunction implies the presence of a distinct neurological syndrome, such as cerebral palsy (Hadders-Algra, 2003). We considered complex MND and major neurological dysfunction as an adverse neurological outcome. Specific types of dysfunction were mild fine motor dysfunction, mild gross motor dysfunction, mild visuomotor dysfunction, or a mild dysfunction in posture and muscle tone. At all ages, children were assessed by K.J.M. under supervision of M.H.-A., who were blind to mode of conception.

Statistical analysis

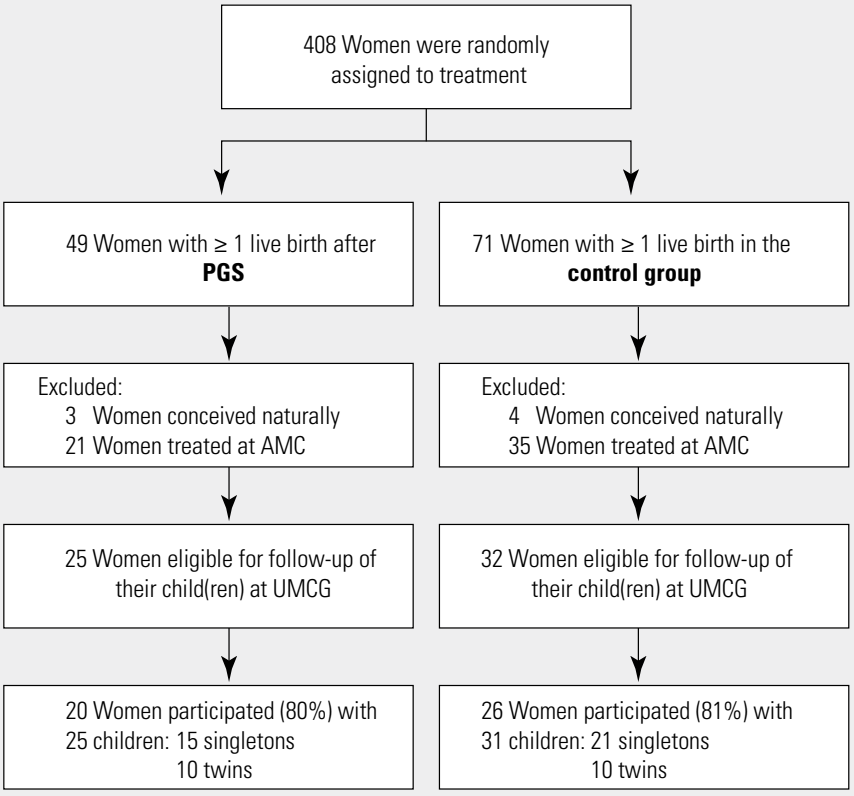
In the preceding study on the efficiency of PGS to establish ongoing pregnancies (Mastenbroek *et al.*, 2007), women were randomised into IVF with or without PGS. In this study, we continued to use women (or their pregnancies) as unit of analysis for the primary statistical analyses. In case of twins, a pregnancy was considered to have an adverse outcome when at least one of the children born had an adverse outcome. For descriptive purposes and secondary analyses, we also analysed data at the child level, under the untested assumption that lack of independence among twins is negligible.

We used t test to analyze differences of means in continuous data. For categorical data, Fisher's exact test or Pearson's χ^2 test was used when appropriate. Because group sizes were small and randomisation had occurred at the level of the women before pregnancy and not on infant level, we *a priori* decided to adjust for factors known to affect neurodevelopmental outcome, i.e. gestational age, twins, and maternal age by means of logistic regression analysis. The p values of 5% and lower were considered statistically significant. Statistical analyses were performed with the use of SPSS for Windows, version 14.0.

RESULTS

Between March 2004 and January 2006, the multicenter trial resulted in live birth of at least one child in 49 women after IVF treatment with PGS and 71 control women (Mastenbroek *et al.*, 2007). Of these women, 25 and 32 gave birth after treatment at the University Medical Center Groningen. Twenty (80%) and 26 (81%) couples agreed to participate in the follow-up study with their child or children (figure 1). Reasons for nonparticipation were logistic (three study and four control couples), unwillingness to participate in a developmental study (two study and one control couples), or resistance against hospital visits due to fertility history (one control couple). The 20 couples of the PGS group participated with 25 children and the 26 couples of the control group with 31 children (table I). Couples with

Figure 1 - Eligibility and participation of women



PGS = Preimplantation genetic screening, UMCG = University Medical Center Groningen, AMC = Academic Medical Center Amsterdam

Table I. Demographic characteristics of parents and children.

Characteristics	Women with ≥ 1 live birth after PGS *(n = 20)*	Children born following PGS *(n = 25)*	Control-women *(n = 26)*	Control-children *(n = 31)*	*P*-values
Parental characteristics:					
Maternal age at conception, mean ± SD	37.6 ± 1.6		38.2 ±1.4		0.24
Primiparity	8 (40)		14 (54)		0.53
Education level mother (high[a]), n (%)	7 (35)		12 (46)		0.65
Education level father (high[a]), n (%)	8 (40)		10 (39)		1.00
Fertility factors:					
Intracytoplasmic sperm injection, n (%)	8 (40)		8 (31)		0.73
Time to pregnancy in years, mean ± SD	3.6 ±1.9		3.9 ±2.8		0.77
Cause of subfertility, n (%)[b]					
poor semen quality	7 (35)		11 (42)		0.84
tubal	8 (40)		9 (35)		0.95
unexplained	4 (20)		7 (27)		0.73
other[c]	3 (15)		3 (12)		1.00
Perinatal characteristics:					
Twins, n (%)	5 (25)		5 (19)		0.73
Gestational age in weeks, mean ± SD	39.1 ± 2.6		38.5 ± 2.0		0.40
Preterm birth (< 37 weeks), n (%)	2 (10)		4 (15)		0.68
Caesarean section, n (%)	3 (15)		7 (27)		0.48
Child chracteristics:					
Male gender, n (%)		15 (60)		15 (48)	0.39
Birth weight in grams, mean ± SD		3187 ± 750		3016 ± 658	0.37

[a] High level of education denotes university education or vocational colleges.
[b] More than one diagnosis per couple was possible
[c] Other causes of subfertility were anovulation, endometriosis and cervical factor.

twin pregnancies participated either with both or none of the children. Postnatal attrition was low. One control child and one pair of study group twins were not assessed at the ages of respectively 2 weeks and 3 and 4 months. All included children were assessed at the ages of 4, 10, and 18 months.

Demographic and perinatal characteristics of the PGS and the control group are shown in table I. Causes of subfertility were similar in the PGS and the control group. Both groups included five twin pregnancies. Testing age (corrected for preterm birth) was similar in study and control group at all follow-up ages.

Five pregnancies after IVF with PGS (25%) compared with four pregnancies in the control group (15%) resulted in the birth of at least one child with an adverse neurological outcome at the age of 18 months (adjusted odds ratio: 2.3 [0.4–12.0]; table II). Neurological outcome at child level is presented in table III. Five PGS children (20%) and four controls (13%) showed complex MND. In addition,

one child in the study group presented with a spastic diplegia. This condition resulted from myelum compression at the level of the medulla oblongata by an arteriovenous malformation diagnosed at the age of 1 year. Analysis of the specific types of dysfunction showed that PGS pregnancies tended to result more often in children with fine motor dysfunction (P = .08) and dysfunctional posture and muscle tone (P = .03; table II). Multivariate analyses were not possible because none of the children in the control group presented with these specific types of dysfunction. There were no differences in gross motor function, reflexes, and visuomotor function between the two groups. Neurodevelopmental outcome up to and including 10 months of age was similar in children born after PGS and control children (data not presented).

Four PGS pregnancies (20%) and three control pregnancies (12%) resulted in admission of at least one child to neonatal intensive care (adjusted odds ratio: 10.2 [0.4–235.2]; table II). Reasons for neonatal intensive care admission were prematurity (four PGS twin children and two control twin children), respiratory insufficiency (two PGS twin children), short-term transitional problems after

Table II - Neurological outcome, neonatal intensive-care admissions and congenital malformations in PGS and control pregnancies.

	Women with ≥ 1 live birth after PGS (n=20)	Control-women (n=26)	Crude Odds-Ratio [95%CI] or p-value	Adjusted[a] Odds-Ratio [95%CI]
Outcome measure	**n (%)**	**n (%)**		
Pregnancies resulting in ≥ 1 child with an adverse neurological outcome at 18 months	5 (25)	4 (15)	1.8 [0.4-8.0]	2.3 [0.4-12.0]
Pregnancies resulting in ≥ 1 child with the following clusters of dysfunction at 18 months				
Fine motor function	3 (15)	0 (0)	P = .08[b]	-
Gross motor function	4 (20)	5 (19)	1.0 [0.2-4.6]	1.0 [0.2-5.0]
Posture and muscle tone	4 (20)	0 (0)	P = .03[b]	-
Reflexes	9 (45)	8 (31)	1.8 [0.5-6.2]	2.0 [0.6-7.6]
Visuomotor function	0 (0)	0 (0)	-	-
Pregnancies resulting in ≥ 1 child admitted to neonatal intensive-care	4 (20)	3 (12)	1.9 [0.4-9.8]	10.2 [0.4-235.2]
Pregnancies resulting in ≥ 1 child with a congenital malformation				
Major malformation	4 (20)	1 (4)	6.2 [0.6-61.1]	4.7 [0.4-51.2]
Minor and/or major malformation	7 (35)	5 (19)	2.3 [0.6-8.6]	2.1 [0.5-9.1]

[a] Adjusted for gestational age, twins, and maternal age at conception.
[b] Fisher's exact test.

TABLE III - NEUROLOGICAL OUTCOME, NEONATAL INTENSIVE-CARE ADMISSIONS AND CONGENITAL MALFORMATIONS IN CHILDREN.

Outcome measure	Children born following PGS (n=25)	Control-children (n=31)
	n (%)	*n (%)*
Neurological outcome at 18 months:		
Normal	18 (72)	26 (84)
Simple MND[a]	1 (4)	1 (3)
Complex MND[a]	5 (20)	4 (13)
Major Neurological Dysfunction	1 (4)[b]	0 (0)
Clusters of dysfunction at 18 months:		
Fine motor function	4 (16)	0 (0)
Gross motor function	5 (20)	5 (16)
Posture and muscle tone	4 (16)	0 (0)
Reflexes	11 (44)	8 (26)
Visuomotor function	0 (0)	0 (0)
Neonatal intensive-care admission	7 (28)	3 (10)
Congenital malformations:		
Children with a major malformation	5 (20)[c]	1 (3)[d]
Children with a malformation (minor and/ or major)	9 (36)	6 (19)

[a] MND = minor neurological dysfunction.
[b] Spastic diplegia due to spinal cord compression by an arteriovenous malformation diagnosed at the age of 1 year.
[c] One child with congenital cataract, one child with a congenital forefoot deformity (metatarsus adductus), one child with an arteriovenous malformation, and a pair of twins with undescended testis requiring orchidopexy.
[d] One child with bilateral inguinal hernia requiring surgery.

birth (one PGS singleton), and sepsis (one control singleton). In total, seven PGS children (28%) and three control children (10%) were admitted to neonatal intensive care (table III). As condition at birth, i.e. the requirement of neonatal intensive care, might be an important mediator of neurological condition at 18 months, we explored this relationship. Only one of the 10 children who had been admitted to neonatal intensive care presented with adverse neurological outcome at 18 months, indicating no statistically significant association (P = .67). In addition, we found no relation between neonatal intensive care admission and the number of dysfunctional clusters present (P = .84) or the presence of any specific type of MND.

In four PGS pregnancies (20%) and one control pregnancy, at least one child presented with a major malformation at the age of 18 months (adjusted odds ratio: 4.7 [0.4–51.2]; table II). As the presence of congenital malformations is associated with an increased risk of neurological dysfunction, we explored this relationship. We found an evident relation between major malformations and adverse neurological outcome at 18 months (P = .007). Children with a major

malformation showed more dysfunctional clusters at the age of 18 months (P < .001) and presented more often with fine motor impairment (P = .05), gross motor impairment (P = .007), and mild dysregulation of posture and muscle tone (P = .003). Congenital malformations of any kind (major or minor) were observed in respectively seven (35%) and five (19%) pregnancies (adjusted odds ratio: 2.1 [0.5–9.1]; table II). These malformations were not related to adverse neurological outcome (P = .11) or the presence of any specific type of MND, however, children with any kind of malformation had more clusters of dysfunctions (P = .04) than children without malformation.

DISCUSSION

In this prospective, assessor-blinded randomised follow-up study, we found similar rates of adverse neurological outcome, neonatal intensive care admission, and congenital malformations in pregnancies resulting from IVF with or without PGS. Mild fine motor dysfunction and mildly dysfunctional posture and muscle tone tended to be present more frequently after PGS pregnancies. At child level, results were slightly more unfavourable for children born after PGS than for control children. Therefore, an increased risk for a less favourable neurological outcome in children born after PGS cannot be excluded.

The main limitation of our study is the relatively small sample size, which was caused by the fact that the power analysis of the PGS study was based on the number of women needed to detect an increase in ongoing pregnancy rates and not on the number of pregnancies necessary for follow-up (Mastenbroek *et al.*, 2007). This means that this study has an explorative nature and that our findings should be interpreted with caution. From an infant developmental perspective, the fact that randomisation had occurred at the level of future mothers and not at infant level might also be regarded as a limitation of the study. We addressed the problem of minor differences in perinatal adversity between the groups by including multivariate statistics. Further minor limitations of the study are the 19–20% loss to follow-up and the relatively limited duration of follow-up. For practical reasons, outcome at the age of 18 months is frequently the end point in developmental studies. However, it should be realised that the majority of minor developmental disabilities first emerge at school age (Hadders-Algra, 2002). Therefore, long-term neurodevelopmental follow-up of children born after IVF with PGS is urgently needed.

The strengths of our study are prospective follow-up, blinding of the assessor, and application of longitudinal detailed, standardised neurological

assessments. Furthermore, randomisation of couples to IVF treatment with or without PGS contributed to the comparability of study and control group.

Many studies have investigated neurodevelopmental outcome of children born after assisted reproduction (reviewed by Sutcliffe and Ludwig, 2007;Middelburg *et al.*, 2008;Hvidtjørn *et al.*, 2009). Results from two recent systematic reviews suggest that IVF is associated with an increased risk for cerebral palsy. At least partly, this association can be explained by the increased risk of multiple births and preterm delivery after IVF (Middelburg *et al.*, 2008;Hvidtjørn *et al.*, 2009). Little is known about the association of assisted reproduction and MND (Middelburg *et al.*, 2008), although these so-called minor neurodevelopmental disorders may have a major impact on daily life of the child and his or her family (Batstra *et al.*, 2003; Gillberg and Gillberg, 1989). Even less is known on outcome of children born after PGS. So far, data on two nonrandomised controlled trials have been reported. Banerjee *et al.* reported similar developmental and behavioural scores in children born after PGS or natural conception (Banerjee *et al.*, 2008). In addition, a Belgian group found no adverse effect of embryo biopsy on growth, congenital malformations, neonatal intensive care admissions, behaviour, and mental and psychomotor development assessed with the Bayley Scales in singletons at 2 years (Nekkebroeck *et al.*, 2008a;Nekkebroeck *et al.*, 2008b;Desmyttere *et al.*, 2008).

Considerations and conclusions

Invasive assisted reproductive techniques such as embryo biopsy carry potential risks for further development of the embryo. Biologically, it is plausible that the use of laser or chemicals for opening of the zona pellucida may induce thermal, mechanical or mutagenic side effects (Kanyo and Konc, 2003). Likewise, intuition suggests that removal of up to one quarter of the embryo's cell mass might have developmental consequences. In addition, a different selection of embryos, on the basis of ploidy-status instead of morphologic criteria of the embryo, might hypothetically result in children with less favourable neurological outcome.

At the RCT level of pregnancies, no adverse effect of PGS could be demonstrated. However, at child level, PGS was associated with a higher rate of MND. Contrary to our expectations, adverse neurological outcome was not associated with neonatal intensive care admission. Adverse neurological outcome was, however, strongly associated with major congenital malformations. Previous studies that found an association between minor congenital malformations and impaired neurodevelopmental outcome, in particular impaired fine motor function, have suggested that the link points to an early ontogenetic origin of the neural dysfunctions (Largo *et al.*, 1989;Soorani-Lunsing *et al.*, 1993). The absence of a relation between neonatal intensive care admission and adverse

neurodevelopmental outcome in this study might be due to the small sample size.

It is important to continue neurological follow-up at least up to school age because children may grow into an adverse neurological outcome. This is a well-known developmental phenomenon. Developmental changes in the infant brain strongly affect the age-specific expression of dysfunction. For instance, early signs of dysfunction in infants may disappear because of restorative mechanisms in the brain, but dysfunction may also be uncovered by developmental progress. Cases in point of the latter are development of dyslexia or the time needed for a full expression of cerebral palsy (Hadders-Algra *et al.*, 2004).

Embryo biopsy is also applied in preimplantation genetic diagnosis; however, the indication for this procedure is fundamentally different from PGS (Sermon *et al.*, 2004). In preimplantation genetic diagnosis, blastomeres obtained with embryo biopsy are analysed for specific heritable disorders, such as cystic fibrosis. There is a steady rise in centres practicing preimplantation genetic diagnosis, and the technique is used for more and more indications (Goossens *et al.*, 2008). The increase in application of embryo biopsy in combination with the findings of this study demonstrates the need for determination of its safety.

In conclusion, in this prospective, assessor-blinded randomised follow-up study, we found similar rates of adverse neurological outcome, neonatal intensive care admission, and congenital malformations in PGS and control pregnancies. Because at child level, rates of adverse outcome were higher in children born after PGS than in control children, an increased risk for a less favourable neurological outcome in children born after PGS cannot be excluded by this study. Biologic plausibility, i.e. the possibility that interference with the very first steps of ontogeny may give rise to impaired neurological development, is reason for concern. The safety of new assisted reproductive techniques, such as PGS, should be carefully evaluated in follow-up studies preceding large-scale implementation of these techniques.

Part III - Follow-up of children born after IVF with Preimplantation Genetic Screening

8

Mental, psychomotor, neurological, and behavioural outcome of 2-year-old children born following preimplantation genetic screening: follow-up of a randomised controlled trial

K.J. Middelburg
M. van der Heide
B. Houtzager
M. Jongbloed-Pereboom
V. Fidler
A.F. Bos
J.H. Kok
M. Hadders-Algra
for the PGS follow-up study group

Fertility and Sterility 2011; 96: 165-169.

ABSTRACT

Objective: To evaluate the effect of preimplantation genetic screening (PGS) on neurodevelopmental outcome in children.
Design: Prospective, assessor-blinded, follow-up study of children born to women randomly assigned to IVF/ICSI with or without PGS.
Setting: University Medical Center, Groningen, and Academic Medical Center, Amsterdam, The Netherlands.
Patients: Fifty-four PGS-children and 77 controls.
Intervention: PGS.
Main outcome measure(s): Mental, psychomotor, neurological and behavioural outcome in two-year-old children measured with the Bayley Scales of Infant Development, the Hempel neurological examination, and the Child Behaviour Check List.
Results: Mental, psychomotor, and behavioural outcome at 2 years in children born following IVF with and without PGS was overall similar. PGS-children showed lower neurological optimality scores than control children. Scores on all tests were within the normal range.
Conclusions: PGS-conception does not seem to be associated with impaired mental, psychomotor, and behavioural outcome at age 2. The lower neurological optimality scores found in PGS-children may signal less favourable long-term neurological outcome in PGS-children. Our findings stress the need of evaluation of safety of new assisted reproductive techniques before large-scale implementation.

INTRODUCTION

In preimplantation genetic screening (PGS), embryos are screened for chromosome aneuploidies with the intent to improve ongoing pregnancy rates after in vitro fertilisation (IVF). In this procedure, the zona pellucida of the embryo is opened with laser or by chemical means so that one or two blastomeres can be removed by biopsy. These blastomeres are then screened for aneuploidy and in this way embryos with an abnormal chromosomal constitution are identified and discarded. Theoretically, selection of the most viable embryos with PGS leads to higher ongoing pregnancy rates (Wilton, 2002). Results of recent randomised controlled trials, however, were unable to confirm this theory (Staessen *et al.*, 2004;Jansen *et al.*, 2008;Schoolcraft *et al.*, 2008;Staessen *et al.*, 2008;Meyer *et al.*, 2009;Debrock *et al.*, 2010). In fact, a reduction in ongoing pregnancies after IVF with PGS has been reported (Mastenbroek *et al.*, 2007;Hardarson *et al.*, 2008).

The invasiveness of PGS necessitates careful follow-up of children born after this procedure. So far, few controlled studies have investigated developmental outcome of PGS-children. Previously, our group reported on developmental outcome during infancy of children born at one centre of a two-centre randomised controlled trial on the effects of PGS. The small study reported no statistically significant differences in neurological outcome between children born after IVF with PGS and those born after IVF without PGS up until the age of eighteen months (Middelburg *et al.*, 2010). However, dysfunction in fine motor abilities and mild dysfunctions in posture and muscle tone regulation tended to be present more frequently following PGS. The small sample size of the study precluded the conclusion that PGS is not associated with a less favourable neurological outcome. So far, only two other groups reported on children born after PGS. They reported similar developmental and behavioural scores in children born following PGS or natural conception (Banerjee *et al.*, 2008) and found no adverse effect of PGS on growth, congenital malformations, neonatal intensive-care admissions, behaviour, and mental and psychomotor development (Nekkebroeck *et al.*, 2008a;Nekkebroeck *et al.*, 2008b;Desmyttere *et al.*, 2008).

The present study evaluates neurodevelopmental outcome of 2 year old PGS-children in comparison with children conceived after conventional IVF. To this aim, we conducted a prospective assessor-blinded follow-up study of children born to women participating in a multicentre randomised controlled trial (RCT) on IVF with or without PGS (Mastenbroek *et al.*, 2007).

METHODOLOGY

Eligibility and recruitment of participants

Eligible for the present follow-up study were all children born to women participating in a double blind RCT on the efficiency of PGS to improve ongoing pregnancy rates after IVF (Mastenbroek *et al.*, 2007). Inclusion criteria for women participating in the trial were: age 35 to 41 years, no previously failed IVF-cycles, and no objections against a possible double embryo transfer. Randomisation of women was performed centrally with minimization for age (35-37 or 38-41 years) and reproductive technique (IVF or ICSI), with stratification according to study centre (University Medical Center Groningen (UMCG) and Academic Medical Center (AMC), Amsterdam) prior to the start of the IVF-procedures. Further details on study design and information concerning IVF-treatment procedures have been reported previously (Mastenbroek *et al.*, 2007;ISRCTN76355836).

Before inclusion and randomisation, couples were informed that a follow-up program was part of the PGS-trial and that children born to couples who were included in the PGS-trial would be invited for neurodevelopmental follow-up.

The protocol of the follow-up study was approved by the Dutch Central Committee of Research Involving Human Subjects and the Medical Ethics Committees of the local hospitals. For the children participating in the follow-up study, parental written informed consent was provided.

Outcome measures

The Dutch standardization of the Bayley Scales of Infant Development - second edition (BSID-II-NL; (Van der Meulen *et al.*, 2004)) was used to evaluate mental and psychomotor development. The mental developmental index (MDI) is determined by performance on items measuring visual and auditory information processing, memory, language development, eye-hand coordination, imitation, and problem-solving abilities. The psychomotor developmental index (PDI) consists of items measuring fine and gross-motor skills. Dutch standardization norms (mean ± standard deviation: 100 ±15) were applied (Van der Meulen *et al.*, 2004). The BSID-II-NL mental and psychomotor scales were administered by trained psychologists.

Neurological function was assessed with the standardised neurological examination according to Hempel (Hempel, 1993), an assessment which pays special attention to the presence of minor neurological dysfunction (MND). In this assessment, fine-motor function, gross-motor function, posture and muscle tone, reflexes, and visuomotor function are tested during standardised play-sessions. Multiple signs in one of these domains result in a so-called dysfunctional domain (Hadders-Algra, 2003). In case of the absence of a dysfunctional domain or only the

presence of dysfunctional reflexes, children are classified as neurologically normal. The isolated presence of fine motor, gross motor, or visuomotor dysfunction or mild dysregulation of posture and muscle tone is defined as simple MND. It reflects a normal, but non-optimal form of brain function (Hadders-Algra, 2003). When two or more domains of dysfunction are present, a child is classified as having complex MND. This type of MND is clinically relevant, since it is associated with learning and behavioural disorders at school age (Hadders-Algra, 2002;Batstra *et al.*, 2003). Distinct neurological syndromes, such as cerebral palsy (CP), are classified as major neurological dysfunction (Hadders-Algra, 2003). We considered complex MND and major neurological dysfunction as an adverse neurological outcome.

In addition, the Hempel examination was used to calculate a neurological optimality score (NOS) and a fluency score by determining the number of items scored in a predefined optimality range (Huisman *et al.*, 1995). Subtle nervous system dysfunction is expressed in reduction of the fluency of movements; therefore the fluency score is sensitive for minimal changes in neuromotor development. Since the range for optimal behaviour is narrower than that for normal behaviour, the NOS is a valuable measure to evaluate subtle differences (Prechtl, 1980). Higher scores on the NOS and fluency scale represent better performance.

Parents were asked to complete the Dutch version of the Child Behaviour Check List (CBCL) 1½ - 5, a questionnaire designed and validated to identify problem behaviour in children aged 1½ to 5 years (Achenbach and Rescorla, 2000). CBCL items are grouped into the following problem scales; emotionally reactive, anxious/depressed, somatic complaints, withdrawn, sleep problems, attention problems and aggressive behaviour. The first four scales together form the internalizing scale (CBCL-int) and the latter two form the externalizing scale (CBCL-ext). The sum of all items determines the total problem scale (CBCL-tot). Higher scores represent more problematic behaviour.

Procedures

Demographic information, including e.g. parity, gestational age, birth weight, neonatal intensive-care unit admission, parental age, and educational level was collected on standardised charts at the follow-up assessments. At both centres, the assessors involved in the study were blinded to mode of conception.

Statistical analyses

We used Student's t-test and Fisher's exact test to compare demographic characteristics of the PGS and control groups. We used mixed-effects linear regression analysis to evaluate the possible effect of PGS on the outcome measures. Continuous outcome variables were transformed by the Box-Cox method in order

to achieve normality of regression residuals. The following transformations were used: (MDI/10)2.5, PDI, -ln(59.5-NOS), -ln(14.5-fluency), ln(CBCL-int), CBCL-ext, and ln(CBCL-tot). Apart from PGS we included the following fixed-effect variables in regression models: gender, gestational age (weeks), twin-status, maternal age at conception, hospital (where the assessments at the age of 2 years took place), birth weight, parity, prematurity status, conception via IVF or ICSI, age at examination and educational level of parents (low, medium or high). As data from the twins are likely to be correlated, the models included also a random effect due to mothers. We have used the results of the mixed-effects analysis to calculate confidence intervals (CI) for adjusted difference between the means of the PGS and control group. To interpret these intervals on the original scale we use the fact that the difference between means of two groups, A and B, on the transformed scale for the NOS score can be interpreted as the logarithm of the ratio (59.5-medB)/(59.5-medA), where medA and medB are medians on the original scale.

To evaluate the effect of PGS on the occurrence of an adverse neurological outcome we used the Fisher's exact test. A correction for potential confounders was not attempted and the dependence of twins was ignored because of a small number of non-normal outcomes.

Statistical analyses were performed using statistical packages R and SPSS 16. P-values of 5% or less were considered significant.

RESULTS

Between March 2004 and January 2006, the trial resulted in 52 ongoing pregnancies after PGS and 74 ongoing control pregnancies (pregnancies following conventional IVF-treatment) (Mastenbroek *et al.*, 2007). Respectively, 49 and 71 pregnancies resulted in the live birth of at least one child (table I). A pair of twins died directly postpartum due to immature birth and four couples withdrew informed consent during IVF-treatments and could therefore not be invited for follow-up. Reason for withdrawal was in most cases psychological stress caused by the blinding of couples to treatment allocation. Subsequently, 47 couples in the PGS group and 68 control couples were eligible for follow-up. Eventually, 45 (96%) PGS-couples with 54 children (36 singletons and 9 twins) and 63 (93%) control couples with 77 children (49 singletons and 14 twins) were included.

Demographic and neonatal characteristics of parents and children are shown in table II. Overall, characteristics of the groups were similar. Exceptions are that mean maternal age at conception was half a year higher in the control group

Table I - Eligibility and participation of couples and their children.

	Preimplantation genetic screening	Control group
Ongoing pregnancies	52	74
Intra-uterine deaths/ terminated pregnancies[a]	3	3
Pregnancies resulting in ≥ 1 live birth	49	71
Pregnancies resulting in postpartum death[b]	1	0
Withdrawal of informed consent[c]	1	3
Elegible couples	47	68
Non-participants[d]	2	5
Participating couples	45	63
with singletons	36	49
with twins	9	14
Children assessed at age 2 years[e]	54	77

[a] In the PGS-group: 1 termination because of trisomy 18, 1 intrauterine death due to abruptio placentae and 1 immature birth. In the control group: 2 terminations because of trisomy 18 and cleft lip and palate and 1 intrauterine death.
[b] postpartum death of twins due to immature birth at 24+5 weeks
[c] Couples whom withdrew informed consent during IVF-treatment were not approached for participation in the follow-up study.
[d] Reasons for non-participation: In the PGS group: 1 moving abroad, 1 assesment burden. In the control group: 1 moving abroad, 2 untraceable, 2 assessment burden.
[e] Analyses according to intention to treat

(P = .04) and gestational age was on average a week shorter in the control group (P = .03) compared to the PGS group.

Table III shows the unadjusted results of the neurological, psychomotor, mental and behavioural assessments. Mixed effects multiple regression analyses revealed no effect of PGS on MDI (P = .468), PDI (P = .540), fluency (P = .074), CBCL-int (P = .635), CBCL-ext (P = .583), and CBCL-tot (P = .442). A significant negative effect of PGS was found on NOS (P = .020). In the analyses the P-values are those from models including all 11 potential confounders. The best fitting models obtained by backward selection of variables included the following variables with a negative contribution to outcome: MDI, CBCL-int, CBCL-ext, and CBCL-tot: lower parental level of education; PDI: male gender and hospital (AMC), NOS: PGS, higher maternal age at conception, hospital (UMCG); and Fluency: higher maternal age at conception, hospital (UMCG). The estimated effect of PGS on -ln(59.5-NOS) was -0.219, the 95% CI was (-0.394 to -0.044). On the original NOS scale this result can be roughly interpreted as a difference of about 2 points (95% CI 0.5 to 4 points) between the corrected median of the PGS group and the corrected median of the control group. The raw unadjusted NOS medians of the two groups were 50 and 51, respectively.

Three children (6%) in the PGS group and one child (1%) in the control group were classified as having an adverse neurological outcome (P = .306, two-

sided). Three of these children had complex MND and one PGS-child was diagnosed with a diplegic cerebral palsy due to a rupture of an arteriovenous malformation at the age of 1 years.

Table II - Demographic characteristics of parents and children.

Characteristics	Couples with ≥ 1 live birth after PGS *(n = 45)*	Children born after PGS *(n = 54)*	Control-couples *(n = 63)*	Control-children *(n = 77)*	*P*-values
Parental characteristics:					
Maternal age at conception, mean (SD)	37.4 (1.4)		37.9 (1.6)		0.040
Primiparity	28 (62%)		46 (73%)		0.294
High education level mother[a]	23 (51%)		32 (51%)		1.000
High education level father[a]	25 (56%)		29 (46%)		0.435
Time to pregnancy in years, mean (SD)[b]	4.3 (2.3)		4.8 (2.9)		0.381
Cause of subfertility[c]:					
poor semen quality	19 (42%)		27 (43%)		
unexplained	13 (29%)		19 (30%)		
tubal	11 (24%)		17 (27%)		
anovulation	2 (4%)		3 (5%)		
endometriosis	2 (4%)		3 (5%)		
cervical	2 (4%)		2 (3%)		
ovarian failure[d]	1 (2%)		0		
Conception method:					0.995
IVF	26 (58%)		36 (57%)		
ICSI	16 (36%)		23 (37%)		
IUI[e]	1 (2%)		1 (2%)		
natural	2 (4%)		3 (5%)		
Perinatal characteristics:					
Twins	9 (20%)		14 (22%)		0.816
Gestational age in days, mean (SD)	275 (16)		268 (18)		0.027
Preterm birth (< 37 weeks)	5 (11%)		11 (18%)		0.420
Caesarean section[b]	8 (18%)		21 (33%)		0.121
Child characteristics:					
Male gender		27 (50%)		42 (54%)	0.722
Birth weight in grams, mean (SD)		3268 (711)		3028 (736)	0.064

[a] high level of education denotes university education or vocational colleges.
[b] 1 missing value.
[c] More than one diagnosis per couple was possible, statistics were not applied
[d] Donated oocytes of women of advanced maternal age were used.
[e] In case of poor follicle growth treatment was converted to intra-uterine insemination. This resulted in birth of a singleton in the PGS group and a twin in the control group.

DISCUSSION

The present study indicated that mental, psychomotor, and behavioural outcome at 2 years in children born following IVF with and without PGS was similar. Median scores on the tests in these domains were well within the normal range for both groups. Yet, PGS-children showed significantly lower neurological optimality scores than control children.

Since the range of optimal behaviour is narrower than the range of normal behaviour, the NOS and fluency score are sensitive measures for minor neurological dysfunction (Prechtl, 1980). The fact that we only found a statistically significant difference between the PGS and control group on the NOS and not on other psychomotor and neurological measures indicates that the neurodevelopmental difference between the two groups is rather subtle. The importance of this finding for long term neurological outcome is currently unknown and therefore continuing follow-up is important to warrant the safety of PGS.

The results of our study are mostly in accordance with the other neurodevelopmental follow-up studies performed so far. Banerjee *et al.* conducted a case-control study in which they compared 49 children born after embryo biopsy - PGS and preimplantation genetic diagnosis (PGD) – to 66 naturally conceived

Table III - Neurological outcome, behaviour and mental and motor perfomance.

Outcome measure	Children born after PGS *(n = 54)*	Control-children *(n = 77)*
Bayley Scales of Infant Development		
MDI, median (p25, p75)	103 (94.5, 110.0)[a]	103 (94.0, 110.5)
PDI, median (p25, p75)	92 (79.0, 103.0)[c]	90 (84.0, 99.0)
Neurological outcome Hempel		
Normal	47 (87%)	73 (95%)
Simple MND	4 (7%)	3 (4%)
Complex MND	2 (4%)	1 (1%)
Cerebral palsy	1 (2%)	0 (0%)
NOS, median (p25, p75)	50.0 (46.75, 52.0)	51.0 (47.0, 54.5)
fluency-score, median (p25, p75)	10.0 (9.0, 11.0)	11.0 (9.5, 11.0)
Child Behaviour Check List		
CBCL-int, median (p25, p75)	41.0 (37.0, 47.0)[a]	43.0 (37.0, 53.0)[b]
CBCL-ext, median (p25, p75)	47.0 (42.5, 56.0)[a]	51.0 (44.0, 56.0)[b]
CBCL-tot, median (p25, p75)	43.0 (38.0, 50.5)[a]	46.0 (41.0, 53.0)[b]

[a] 1 missing value
[b] 2 missing values
[c] 3 missing values

control children (Banerjee *et al.*, 2008). Remarkably, PGS/PGD-children scored significantly lower on the loco-motor subscale of the Griffiths than the controls, while the Griffiths General quotient and the results of the Toddler Temperament Questionnaire were similar in both groups. Nekkebroeck *et al.* compared a cohort of PGS/PGD-children with cohorts of ICSI and naturally conceived (NC) children (Nekkebroeck *et al.*, 2008a;Nekkebroeck *et al.*, 2008b). Consistent with our findings, they found no between-group differences for Bayley's MDI and PDI (Nekkebroeck *et al.*, 2008a). In addition, they found that PGS/PGD and ICSI mothers reported fewer CBCL total problems than NC mothers, whereas CBCL externalizing problems were reported fewer by ICSI mothers than PGS/PGD and NC mothers. ICSI fathers reported fewer total and external problems than PGS and NC fathers (Nekkebroeck *et al.*, 2008b).

Strengths of our study are its design, blinding of the assessors, minimal attrition, extensive correction for confounders, and the use of sensitive and validated neurodevelopmental outcome measures. This study is the first follow-up study of children born after IVF in which couples were randomised into IVF with or without PGS. Couples were informed on the follow-up when they were included in the trial (before start of the IVF-treatments), which resulted in high participation rates. Since twins as well as singletons were included, data are representative for children born following PGS.

A limitation of the present study is the fact that the power analysis was based on the number of women needed to detect an increase in ongoing pregnancy rate (Mastenbroek *et al.*, 2007) instead of the number of children needed for follow-up. The unforeseen negative effect of PGS on ongoing pregnancies further reduced the number of children available for follow-up. Future meta-analysis would be helpful to overcome sample size limitations of current follow-up studies. Uniformity of neurodevelopmental assessments is crucial for this matter.

Even though PGS, as practiced in the present trial, does not improve live birth rates, follow-up of children born after embryo-biopsy is still of importance. A recent meta-analysis confirmed that ongoing pregnancy rates following PGS are reduced (Mastenbroek *et al.*, 2008). Therefore the technique is no longer practiced on routine basis. Currently, it is studied whether alternative forms of PGS (such as polar body biopsy and more comprehensive chromosome testing) are feasible and efficient in improving live birth rates (Harper *et al.*, 2010;Geraedts *et al.*, 2010). Embryo-biopsy is also applied in PGD, a technique in which blastomeres are tested for specific heritable disorders in at-risk patients (Sermon *et al.*, 2004). Continuation of the use of embryo-biopsy justifies follow-up and reassessment of the children in the current trial at later ages may yield valuable information concerning the effects of embryo-biopsy on later development and health.

In conclusion, the results of the present follow-up study showed similar outcome in terms of mental, psychomotor and behavioural outcome in children born following PGS and conventional IVF. Nevertheless, PGS-children scored lower on the NOS, a marker for subtle neurological dysfunction. The importance of the latter finding for long term neurological outcome is currently unknown. Continuation of follow-up is needed to warrant the safety of PGS.

Part IV
General discussion and future perspectives

9

This thesis describes the neurodevelopmental outcome of children born following assisted reproductive technology. Series of sensitive, validated measures have been used to scrutinize neurodevelopment up to the age of two years. Overall, results of the studies in this thesis have been reassuring. In the Groningen ART cohort study, we found no differences in neurodevelopmental outcome - including mental, psychomotor, neurological and behavioural development - of children born following ART with and without ovarian hyperstimulation up to the age of two years. Additionally, follow-up of children born after ART with preimplantation genetic screening showed no association between PGS conception and impaired mental, psychomotor, and behavioural outcome at age two; however an association was found between PGS conception and lower neurological optimality scores.

The human central nervous system continues to develop after birth. Large part of myelination and synapse formation and elimination occurs throughout childhood and adolescence (Graaf-Peters and Hadders-Algra, 2006). Therefore, the age of two years is too early to draw pertinent conclusions on neurodevelopmental outcome of ART children. Continuation of follow-up and meta-analysis of follow-up studies is still needed to warrant the safety of assisted reproductive technology on long term.

Methodological considerations

True distinction between the effects of ART and the underlying indication for treatment can only be made in a trial with random allocation of assisted and natural conception, which is unethical (Buck Louis *et al.*, 2005;Knoester, 2007). In practice, the possibilities for follow-up studies are therefore limited. Best evidence comes from register based studies and prospective cohort studies. The former have proven to be very valuable when studying disorders of low prevalence, such as cerebral palsy (Ericson *et al.*, 2002;Stromberg *et al.*, 2002;Pinborg *et al.*, 2004;Klemetti *et al.*, 2006). However, register based studies do not allow for detailed study of mental, psychomotor, neurological and behavioural development. For this matter we need controlled studies. The systematic review taught us that the number of controlled studies with robust methodological quality is still limited, since follow-up studies are often hampered by practical difficulties. In the following paragraphs, the methodology of the studies described in this thesis will be addressed.

The Groningen ART-cohort study

This prospective, longitudinal cohort-study is unique for its study groups; children born after IVF/ICSI with controlled ovarian hyperstimulation (COH-IVF) are compared to children born after IVF/ICSI in a modified natural cycle (MNC-IVF), and a control group consisting of naturally conceived children born to couples who were waiting for subfertility work-up. With this design we aimed at disentangling the potential effects of controlled ovarian hyperstimulation and the *in vitro* procedure itself on neurodevelopmental outcome. Nevertheless, the medication used in MNC-IVF - although minimal - may have caused an overestimation of the effect of IVF or an underestimation of the effect of COH. In the interpretation of the results of our study we were not hampered by this minor confounding of MNC, since assisted reproduction was not associated with reduced neurodevelopmental outcome in any of the studies. An alternative method to analyze the effect of ovarian stimulation on neurodevelopmental outcome would be to compare outcome of naturally conceived children to children born following ovulation induction. A drawback of this method is that the latter procedure aims at obtaining only 1 or 2 follicles, while in COH-IVF the aim is to harvest up to 10 oocytes (Nargund *et al.*, 2007). Thus, studying children born following ovulation induction is not truly representative of the effects of ovarian hyperstimulation.

Another unique aspect of this study is its exceptionally high follow-up rate (97-98% at age two). Utmost effort was put into maintaining children in follow-up. Parents were involved in the study by means of regular examinations since birth and newsletters concerning the progress of the study. Furthermore, home-visits were made when parents were unable to visit the hospital. But most importantly, the children enjoyed the neurodevelopmental examinations, as these are very playful. Despite the high follow-up rate, sample size of the study is still relatively small, which is a limitation of the study. It precludes any definite conclusion on disorders of low incidence.

It may be assumed that the low attrition, together with good initial prenatal enrolment (63-76%) has yielded a representative sample of singleton ART children in Groningen. Generalisability to all ART children is limited, since we excluded twins. COH-IVF is associated with an increased prevalence of twins. We excluded twins from the analyses as we were unable to identify a sufficiently large group of twins born following MNC-IVF and naturally conceived twins born to subfertile patients. Selection bias was studied by comparing characteristics of participants and non-participants. Participation was largely non-selective, with the exception of maternal age in NC mothers (non-participating mothers were younger).

We applied detailed, standardised neurodevelopmental assessments at the corrected ages of 2 weeks, 3, 4, 10, 18 and 24 months. The instruments used allowed us to study subtle differences in neurodevelopmental outcome early in life. The advantage of assessing neurodevelopmental outcome at early age is the relatively limited influence of postnatal factors, such as social conditions, on developmental outcome. This allows for a closer linkage of findings to early ontogenetic events. Data on the long term predictive value of the neurodevelopmental instruments used in this thesis (GMs, TINE and Hempel) in groups of infants not considered at high risk for developmental disorder is scarce. However, reliability measured by inter- and intra-assessor agreement, of the instruments is good (Hempel, 1993;Hadders-Algra *et al.*, 2009;Bouwstra *et al.*, 2009;Middelburg *et al.*, 2010). Therefore the instruments can be used to compare groups; however, it is unknown what the exact significance of the current findings is for the children's long-term neurodevelopmental outcome.

Notably, the secondary analyses (the comparison between the subfertile group and fertile reference group) in the GM study and the studies concerning developmental follow up at age two are hampered by the fact that the fertile reference groups were recruited separately and postnatally. Another limitation of these comparisons was that the assessors were not blind to group allocation, which reduced internal validity, and with that, external validity of the comparison. In addition, for the reference infants in the GM-study no information was available on conception method. It is rather likely that some of the children were born following assisted reproduction. This may imply that actual differences in GM-quality between children of fertile and subfertile couples were larger than indicated in the study.

Follow-up of children born after IVF with Preimplantation Genetic Screening

To our knowledge this study is the first prospective follow-up study of children born after IVF in which couples were randomised into IVF with or without PGS. Randomisation contributed to the comparability of the study and control group, it resulted in similar demographic characteristics in the two groups.

The main limitation of this study is its sample size. The power analysis was based on the number of women needed to detect an increase in ongoing pregnancy rate instead of children needed for follow-up. Unfortunately, the number of children available for follow-up was limited by the negative effect of PGS on ongoing pregnancies (Mastenbroek *et al.*, 2007). In the study reporting on neurodevelopmental outcome from 2 weeks to 18 months, sample size was even further reduced. Due to logistical reasons, only children living in the northern part of the Netherlands were studied at these ages.

Since both twins and singletons were included in this study, data on neurodevelopmental outcome are representative for all children born following PGS. Generalisability is also supported by the relatively high participation rates (~80%), which may likely be a result of the fact that couples were informed on the follow-up before start of the IVF-treatments. Prenatal or - even better - preconceptional inclusion of children helps to prevent selection bias, as later inclusion may be affected by the child's neurodevelopmental status. Parents of children with a neurodevelopmental disorder may be inclined to refrain from participation in follow-up research for two reasons. First, confrontation with the limitations of their child during assessment. Secondly, these parents frequently visit the hospital for follow-up of their child's disorder and therefore time constraints may play a role. On the other hand, parents who are concerned about their child's development may be eager to participate, with the intention of receiving a professional opinion on the child's development. Similarly, parents of children with advanced development are often willing to cooperate in our experience.

The influence of ovarian hyperstimulation and ART laboratory procedures on neurodevelopmental outcome

Factors that may affect perinatal outcome after ART may also influence development and health of children born after ART. The mechanisms underlying these relations are as yet unclear. Hypothetically, embryo development may be affected by one or more components of the ART-procedure. For instance, several natural selection procedures are bypassed in ART and culture conditions may cause disturbed genomic imprinting (Ceelen and Vermeiden, 2001;Young *et al.*, 2001;Khosla *et al.*, 2001). Furthermore, altered hormone levels in ART may interfere with nidation and placentation, leading to suboptimal uterine conditions (Ertzeid and Storeng, 2001;Ceelen and Vermeiden, 2001;van der Auwera and D'Hooghe, 2001). Research so far primarily focussed on perinatal outcome measures, such as birth weight and gestational age. Nevertheless, it is possible that consequences extend beyond the perinatal period (Barker, 1995).

In our studies, we chose to correct for birthweight and gestational age by means of multivariate statistics as our research question was whether ovarian hyperstimulation or the in vitro procedure affected outcome at 2 years, given the potential effect of assisted reproduction on perinatal outcome. One could also argue not to correct for these factors, since they are mediators on the pathway from assisted reproduction to neurodevelopmental outcome. Therefore we repeated the analysis without the mediating factors birth weight and gestational age. Overall, this did not affect the results (reported in Chapters 4 and 6).

The results of the systematic review showed no consistent differences in neuromotor, cognitive, language en behavioural development between children born after ART and naturally conceived children. Evidence is sufficient to conclude that gross developmental pathology in children born following ART is absent up to the age of three years. Specific gross effects of ovarian hyperstimulation or the ART laboratory procedures are therefore neither expected in this age period. Nonetheless, the measures used to assess developmental outcome in most studies (e.g. the Bayley Scales of infant Development) are not capable of evaluation of neuromotor and cognitive functioning in a detailed sense. Subtle deviations may have little clinical relevance for individual functioning; however subtle deviations in a substantial subpopulation may have impact on society at large. For this reason, it is justified to scrutinise neurodevelopment of the continuously increasing percentage of children born following assisted reproduction.

The relation between subfertility and neurodevelopmental outcome: cause or consequence?

Subfertility itself may also contribute to outcome of ART children. It is known that subfertile couples have an increased risk of obstetric complications and adverse perinatal outcome, such as increased risk of preeclampsia, antepartum haemorrhage, caesarean section, preterm birth, low birth weight and perinatal death (Draper *et al.*, 1999;Pandian *et al.*, 2001;Thomson *et al.*, 2005).

In the Groningen ART-cohort, we found a remarkable high percentage of the naturally conceived pregnancies of subfertile couples to be complicated by pregnancy induced hypertension. Furthermore, the naturally conceived children of subfertile couples were relatively often born by caesarean section, showed more signs of foetal distress, and were more frequently admitted to neonatal intensive care than children of the ART-groups, even though birth weight and gestational age at birth were higher. Possibly, this was a chance-finding. Our study was not designed, nor powered for these outcome measures. But theoretically, these findings may also have been the result of an increased risk in pregnancies of subfertile couples in combination with less intensive obstetrical care than provided in ART-pregnancies. If the association between subfertility and perinatal outcome is clarified, this may help to provide customised obstetrical care to patients at risk (Thomson *et al.*, 2005).

At the age of three months, we saw a slightly reduced general movement quality in children born to subfertile couples compared to a reference group. Whereas, at the age of two years, we found better neurological outcome in the

subfertile group than in another reference group. The inconsistency of these findings may be the result of selection bias. The reference group at 3 months was representative for the general population. Children were recruited and assessed as part of a general health check-up provided to all Dutch children. At the age of two years, parents were invited to volunteer at the child-welfare-centre, but a separate appointment was made for the examinations. The total time scheduled for the assessments was nearly two hours per child. The effort that had to be taken may have led to selection of parents with concerns about their child's development.

The relation between subfertility and neurodevelopmental outcome was also studied in the Danish national cohort study. Sun *et al.* reported no increased prevalence of epilepsy and febrile seizures in children of untreated subfertile couples compared to children of fertile couples (Sun *et al.*, 2007) and Zhu *et al.* did not find an association between time to pregnancy and age of milestone achievement or risk for cerebral palsy (Zhu *et al.*, 2009;Zhu *et al.*, 2010a). Yet, Zhu et al did find a modestly increased risk of developmental coordination disorder (Zhu *et al.*, 2010b). Altogether, solid evidence of an association between subfertility and neurodevelopment is absent. Nevertheless, it is conceivable that a non-optimal genetic make-up or hormonal condition results not only in subfertility, but also in less optimal neurodevelopmental outcome in offspring.

Preimplantation genetic screening and neurodevelopmental outcome

Theoretically, embryo biopsy as performed in PGS may induce damage which interferes with the embryo's further development. For instance, the use of laser or chemicals for opening the zona pellucida may induce thermal, mechanical or mutagenic side effects (Kanyo and Konc, 2003) with long-lasting consequences.

In the PGS-study we found a statistically significant difference between the PGS and control group on the neurological optimality score (NOS) at 2 years, but not on other psychomotor and neurological measures. This finding indicates that the difference found is rather subtle, since the range of optimal behaviour is narrower than the range of normal behaviour (Prechtl, 1980). Other groups who studied PGS-children reported similar outcome in study and control groups, except for lower scores on the loco-motor subscale of the Griffiths in children born after PGS (Banerjee *et al.*, 2008;Nekkebroeck *et al.*, 2008a;Nekkebroeck *et al.*, 2008b). Altogether, an increased risk for a less favourable neurological outcome after PGS cannot be excluded by evidence available.

Suggestions for future research

Neurodevelopmental disorders may emerge when children grow older and cognitive functioning evolves. The body of evidence of good methodological quality concerning neurodevelopment of children born after ART beyond pre-school age is still limited, therefore this should continue to be a focus of follow-up research. Future meta-analyses would be helpful to overcome sample size limitations in current follow-up studies. Crucial in this matter is uniformity of neurodevelopmental assessments. Louise Brown - the first and therefore eldest person born after IVF - has currently only reached her mid-thirties (Steptoe and Edwards, 1978). Thus, long-term or trans-generational consequences of assisted reproduction may still appear when the cohort of people born after ART grows older. In this aspect, reproductive functioning of people conceived with ART also deserves attention.

The most important factor influencing neurodevelopmental outcome after ART is twin-status. Twin pregnancies carry an increased obstetrical risk and therefore medical specialists consider them less desirable (Land and Evers, 2004). The majority of subfertile patients, however, prefer twins over singletons (van Wely *et al.*, 2006). Understandably, parents often focus on the higher pregnancy rates after double embryo transfer. Additional risks in terms of neurodevelopmental outcome of ART-twins compared to ART-singletons are only known to a limited extent. Information available is largely based on nation-wide register based studies. Remarkably, the reported prevalence of cerebral palsy in ART-singletons and ART-twins is not significantly different, which may be attributable to an increased risk in ART-singletons due to higher prematurity and low birthweight rates (Pinborg, 2005). Controlled studies showed that ART-twins had significantly lower scores on cognitive functioning than ART-singletons (Bonduelle *et al.*, 2003;Olivennes *et al.*, 2005). Information on minor neurological dysfunctioning of ART-twins is not available. The information may help caregivers to counsel patients on the risks of twins and, with that, the risks of treatment strategies that may result in birth of twins, such as ART with controlled ovarian hyperstimulation and double embryo transfer.

The influence of subfertility *per se* on developmental outcome in children may be further investigated by studying children of couples with different indications for assisted reproduction. For instance, tubal pathology is expected to have less effect than unexplained subfertility or subfertility due to male factor. Similarly, time to pregnancy may be related to developmental outcome.

Recently, a growing interest in cardiovascular outcomes following ART arose (Painter and Roseboom, 2007). It is known that the early environment of developing

organisms is important in determining later cardiovascular health (Barker, 1995). In animals, assisted reproduction results, for instance, in hypertension in mice and an increased cardiovascular risk in cattle (Rerat *et al.*, 2005;Watkins *et al.*, 2007). Moreover, preliminary evidence is provided that ART children are also more prone to obesity, hypertension and diabetes (Belva *et al.*, 2007;Ceelen *et al.*, 2008). These findings deserve further attention in future follow-up studies.

Technological possibilities in assisted reproduction develop continuously. New methods are often introduced in practice without a pre-clinical phase, as results in animal studies cannot easily be extrapolated to humans. Moreover, clinical studies in healthy volunteers are also impossible. Therefore, the only way to evaluate the safety of new assisted reproductive techniques is in retrospect. As yet, the knowledge on development of children born following preimplantation genetic screening, preimplantation genetic diagnosis, cryopreservation and in vitro maturation is limited at best. It is of utmost importance that the safety for offspring is evaluated before large scale implementation of these methods.

Concluding remarks

Results of the studies in this thesis have been reassuring. We found no differences in neurodevelopmental outcome - including mental, psychomotor, neurological and behavioural development - up to the age of two years in children born following ART with and without ovarian hyperstimulation and a naturally conceived control group born to subfertile parents. The follow-up study of children born after IVF with PGS showed similar mental, psychomotor and behavioural outcome in 2-year-old PGS and control children. Neurological optimality scores were lower in PGS-children. Long term follow-up is still needed to warrant safety of ART as neurodevelopmental disorders may emerge when children grow older.

10

Summary & Samenvatting

SUMMARY

Chapter 1. General introduction

The number of children born following assisted reproductive technology (ART) has shown a steady rise during the last decades. As a consequence of the growing number of children born after ART, health and development of these children has become of general importance. In this thesis the neurodevelopmental outcome of children born after ART is evaluated up to the age of two years. An attempt is made to unravel the biological mechanisms that may underlie potentially poorer neurodevelopmental outcome. The thesis is divided into 4 parts.

Part I; Overview of the literature

Chapter 2 contains a systematic review of studies which compare children born following IVF/ICSI to children born after natural conception by assessing outcome in terms of neuromotor development, cognition, speech/language and behaviour. Specific attention is paid to the studies' methodological quality based on study design, attrition, blinding of the assessor, validity of the neurodevelopmental tests used, confounders included and group size or power analysis. Twenty-three out of 59 studies had a good methodological quality including 9 register-based and 14 controlled studies. The register-based studies suggested that IVF/ICSI *per se* does not increase the risk for severe cognitive impairment (i.e. mental retardation) or neuromotor handicaps such as cerebral palsy (CP), the association of IVF/ICSI and CP being brought about by the association of assisted conception with risk factors, like preterm birth. In general, controlled studies of good quality did not report an excess of neurodevelopmental disorders in IVF/ICSI-children. However, the majority of studies followed the children during infancy only, thereby precluding pertinent conclusions on the risk of neurodevelopmental disorders that come to the expression at older ages, such as fine manipulative disability or dyslexia. Currently, a negative effect of assisted conception on the developing human brain has not been identified; however, further research of high methodological quality in children beyond pre-school age is needed.

Part II: The Groningen ART-cohort study

We investigated the effect of ovarian hyperstimulation and the in vitro procedure on neurodevelopmental outcome in cohorts of children born following conventional COH-IVF, MNC-IVF and NC children born to subfertile parents.

In **chapter 3**, we evaluated specific effects of ovarian hyperstimulation, the in vitro procedure, and a history of subfertility on neuromotor development at 3 months of age. Participants were 68 singletons born after COH-IVF, 57 singletons

born after MNC-IVF, and 90 NC singletons born to subfertile couples. To study the effect of subfertility itself, data from a fertile reference population (n = 450) were used. Early neuromotor development was measured by means of the quality of General Movements at 2 weeks and 3 months. Quality of general movements (GMs) may be classified as normal-optimal, normal-suboptimal, mildly abnormal, or definitely abnormal. Definitely abnormal GMs indicate brain dysfunction, mildly abnormal GMs normal but non-optimal brain function. Mildly abnormal and definitely abnormal GMs were observed equally frequently in COH-IVF, MNC-IVF, and NC singletons. But the three subfertile groups showed more mildly abnormal GMs than the reference population. This suggests that factors associated with subfertility rather than those related to IVF procedures may be associated with less-optimal early neurodevelopmental outcome.

Chapter 4 documents the neurological condition of the Groningen ART-cohort children measured with the Touwen Infant Neurological Investigation (TINE) at 4 and 10 months and the Hempel examination at 18 months. Both measures, TINE and Hempel, focus on the occurrence of minor neurological dysfunction. Due to the growing number of children born following assisted reproduction technology, even subtle changes in the children's health and development are of importance to society at large. Neurological examination resulted in a neurological optimality score (NOS), a fluency score and a clinical neurological classification. Fluency of movements is easily affected by neurological dysfunction and is therefore a sensitive measure for minimal changes in neuromotor development. The NOS and the fluency score were similar in COH-IVF, MNC-IVF and NC children. None of the children showed major neurological dysfunction and rates of minor neurological dysfunction at the three ages were not different between the three conception groups. In conclusion, we found no effects of ovarian hyperstimulation or the in vitro procedure itself on neurological outcome in children aged 4-18 months.

Chapter 5 evaluates the neurological condition of the Groningen ART-cohort children with the Hempel examination at 2 years. Besides the effects of ovarian hyperstimulation and the IVF laboratory procedures, the effect of a history of subfertility was studied by comparison of the cohort to a reference group of children born to fertile couples (n=101). The fluency score, NOS and the prevalence of MND were similar in COH-IVF, MNC-IVF and NC children. However, the fluency score and NOS of the three subfertile groups were higher and the prevalence of MND was lower compared to the reference group. In conclusion, neurological condition of two-year-olds born after ART is similar to that of children of subfertile couples conceived naturally. Moreover, subfertility does not seem to be associated with a worse neurological outcome at the age of two years.

In **Chapter 6**, we studied the mental, psychomotor and behavioural development of the Groningen ART-cohort children with the Bayley Scales of Infant Development and Achenbach's Child Behaviour Checklist (CBCL). Again, the effect of a history of subfertility was studied by comparison to the reference group of children born to fertile couples. Cognitive, psychomotor and behavioural outcome in COH-IVF, MNC-IVF and NC groups was similar. Developmental outcome and behaviour of the subfertile groups was largely similar to that of the fertile reference group. Nevertheless, the subfertile groups scored higher on the scale of anxious-depressed behaviour than the reference group. In conclusion, the study suggests that cognitive and psychomotor development and behaviour at 2 years are not affected by ovarian hyperstimulation, the in vitro procedure or a history of subfertility.

Altogether the results of the studies of the Groningen ART-cohort are reassuring. It should be kept in mind that subtle neurodevelopmental disorders may emerge when children grow older. Continuation of follow-up is therefore still needed.

Part III; Follow-up of children born after IVF with Preimplantation Genetic Screening

This part evaluates the effect of preimplantation genetic screening (PGS) on neurodevelopmental outcome in children.

In **chapter 7**, we conducted a prospective follow-up study of children born to women randomly assigned to in-vitro fertilisation with or without PGS. Primary outcome was adverse neurological outcome at 18 months; secondary outcome were types of minor neurological dysfunction, neurological outcome before 18 months, neonatal intensive-care admission, and congenital malformations. Twenty women in the PGS-group participated with 25 children and 26 women in the control group participated with 31 children. Five PGS-pregnancies (25%) and 4 control-pregnancies (15%) resulted in birth of at least one child with an adverse neurological outcome. Dysfunction in fine-motor abilities and posture-and-muscle-tone dysregulation tended to be present more frequently following PGS. Neurological outcome before 18 months, neonatal intensive-care admission and prevalence of congenital malformations were similar in study and control pregnancies. Nevertheless, at child level, rates of adverse outcome were higher following PGS. In conclusion, outcome in pregnancies following IVF with and without PGS was similar. The small sample size precludes the conclusion that PGS is not associated with less favourable neurological outcome.

Chapter 8 describes the results of a follow-up study of children born to women randomly assigned to IVF with or without PGS. Mental, psychomotor,

neurological and behavioural outcome in two-year-old children was measured with the Bayley Scales of Infant Development, the Hempel neurological examination, and the Child Behaviour Check List. Participants were 54 PGS-children and 77 controls. Mental, psychomotor, and behavioural outcome at 2 years in children born following IVF with and without PGS was overall similar. Children born after PGS showed lower neurological optimality scores than control children. Scores on all tests were within the normal range. In conclusion, PGS-conception does not seem to be associated with decreased mental, psychomotor, and behavioural outcome at age 2. The lower neurological optimality scores found in children born after PGS may signal less favourable long-term neurological outcome after application of PGS.

The findings in the follow-up studies of children born after PGS stress the need of evaluation of safety of new assisted reproductive techniques before large-scale implementation.

Part IV: General discussion and future perspectives

In **Chapter 9** the findings of the studies are discussed. The results of the Groningen ART cohort study suggest no effect of ovarian hyperstimulation or the IVF procedure itself on neurodevelopmental outcome until the age of 2 years. Furthermore, follow-up of children born after ART with PGS showed no association between PGS conception and decreased mental, psychomotor, and behavioural outcome at age two, however an association was found between PGS conception and lower neurological optimality scores.

Strengths of the Groningen ART cohort study are the unique study groups, the high follow-up rate, the representative participants and the detailed, standardised, longitudinal assessments. Limitations are the relatively small sample sizes and the relatively young age at follow-up. The PGS-study's strengths are the randomisation of subfertile couples, and high participation rates. The main weakness is the relatively small sample size.

Subfertility is associated with adverse perinatal outcome. The relation between subfertility and neurodevelopmental outcome was - just as in other reports - inconsistent in our studies, possibly due to selection bias. Elucidation of the effects of subfertility would be helpful to customise obstetrical and child-welfare-care.

Future research should focus on neurodevelopmental outcome of older children, adolescents and twins. Furthermore cardiovascular outcome of ART-children deserves attention. The safety of new assisted reproductive technologies should be evaluated before large-scale implementation.

SAMENVATTING

Hoofdstuk 1. Inleiding

Het aantal kinderen dat wordt geboren na hulp bij voortplanting is het laatste decennium aanzienlijk gestegen. De gezondheid en ontwikkeling van deze kinderen is hierdoor van algemeen belang geworden. In dit proefschrift wordt de neurologische, motorische en cognitieve ontwikkeling van kinderen geboren na geassisteerde voortplantingstechnieken tot een leeftijd van 2 jaar geëvalueerd. Onderliggende biologische mechanismen die een potentieel slechtere ontwikkeling zouden kunnen verklaren worden bestudeerd. Het proefschrift bestaat uit 4 delen.

Deel I; Literatuuroverzicht

Hoofdstuk 2 bevat een systematisch overzichtsartikel van studies die de ontwikkeling van kinderen geboren na *in vitro* fertilisatie (IVF) of intracytoplasmatische sperma injectie (ICSI) vergelijken met de ontwikkeling van kinderen die na een natuurlijke conceptie zijn ontstaan. Als uitkomstmaten voor ontwikkeling zijn neuromotore ontwikkeling, cognitie, taalontwikkeling en gedrag beschreven. De methodologische kwaliteit van de studies is beoordeeld aan de hand van de onderzoeksopzet, uitval gedurende de follow-up, de blindering van de onderzoekers, de validiteit van de gebruikte onderzoeksinstrumenten, de confounders die in de analyse werden meegenomen en de grootte van de onderzoeksgroep of de aanwezigheid van een power analyse. Drieëntwintig van 59 studies zijn van goede kwaliteit bevonden, waarvan 9 studies gebaseerd op nationale registers en 14 cohortstudies. De registerstudies suggereren dat IVF/ICSI op zichzelf de kans op ernstige cognitieve of neuromotore stoornissen niet vergroot, maar dat de wel aanwezige associatie tussen IVF/ICSI en cerebrale parese wordt veroorzaakt door een associatie tussen voortplantingstechnieken en risicofactoren voor neuromotore stoornissen, zoals vroeggeboorte. De cohortstudies van goede kwaliteit rapporteerden over het algemeen geen overmaat aan ontwikkelingsstoornissen na IVF of ICSI. Echter, het overgrote deel van de studies onderzocht kinderen tot de peuterleeftijd en kan daardoor geen uitspraak doen over stoornissen die op oudere leeftijd ontstaan (zoals stoornissen in de fijne motoriek of dyslexie). Samenvattend is er tot nu toe geen negatief effect van voortplantingstechnieken op de neurologische ontwikkeling vastgesteld, maar kwalitatief hoogwaardig onderzoek met kinderen vanaf schoolleeftijd is schaars.

Deel II: Het Groningse cohort van kinderen geboren na voortplantingstechnieken

Het Groningse onderzoek van kinderen geboren na voortplantingstechnieken evalueert de afzonderlijke effecten van gecontroleerde ovariële hyperstimulatie

(COH) en de IVF procedure op de neurologische ontwikkeling. Groepen kinderen geboren na conventionele IVF, dan wel IVF in de gemodificeerde natuurlijke cyclus (MNC-IVF), en kinderen van subfertiele ouders na een natuurlijke conceptie (NC) worden hiertoe met elkaar vergeleken.

In **hoofdstuk 3** wordt de neuromotore ontwikkeling van kinderen in bovengenoemde groepen vergeleken op de leeftijd van 3 maanden. Achtenzestig eenlingen geboren na COH-IVF, 57 eenlingen geboren na MNC-IVF, en 90 NC eenlingen werden vergeleken. Aanvullend werd het effect van subfertiliteit bestudeerd door de data te vergelijken met gegevens van de algemene, grotendeels fertiele referentiepopulatie (n = 450). De vroege neuromotore ontwikkeling kan worden gemeten met behulp van de kwaliteit van gegeneraliseerde bewegingen op de leeftijden van 2 weken en 3 maanden. Licht en duidelijk afwijkende bewegingspatronen werden even vaak geobserveerd in eenlingen geboren na COH-IVF, MNC-IVF en een natuurlijke conceptie. Echter de eenlingen in deze drie subfertiele groepen lieten vaker licht afwijkende bewegingen zien dan eenlingen in de referentie populatie. Door deze bevindingen lijkt een minder optimale ontwikkeling eerder geassocieerd met subfertiliteit dan met voortplantingstechnieken.

In **hoofdstuk 4** wordt de neurologische conditie van de kinderen in het cohort op de leeftijden van 4, 10 en 18 maanden onderzocht. De nadruk ligt hier op het voorkomen van lichte neurologische disfunctie. Door het groeiende aantal kinderen dat momenteel ontstaat met behulp van voortplantingstechnieken, worden ook subtiele ontwikkelingsstoornissen en gezondheidsproblemen van belang voor de samenleving. Het neurologisch onderzoek resulteert in een neurologische optimaliteitsscore (NOS), een score over het vloeiende bewegingsverloop en een klinische neurologische classificatie. Minimaal neurologisch disfunctioneren leidt al snel tot minder vloeiend verlopende bewegingen en is daardoor een sensitieve maat voor lichte ontwikkelingsproblematiek. Er werd in dit onderzoek geen verschil gevonden in de NOS, het vloeiend verlopen van de bewegingen, of het voorkomen van lichte neurologische disfunctie van eenlingen na COH-IVF, MNC-IVF of eenlingen na een natuurlijke conceptie. Er werd in geen van de groepen ernstige neurologische afwijkingen geconstateerd. Samenvattend werd er geen effect van ovariële hyperstimulatie of de IVF procedure op de neurologische uitkomst van kinderen van 4 tot 18 maanden aangetoond.

In **hoofdstuk 5** wordt de neurologische conditie van de kinderen op de leeftijd van 2 jaar geëvalueerd. Er werd opnieuw gebruik gemaakt van data van een referentiegroep, dit maal een geselecteerde fertiele referentiegroep (n=101), om ook het effect van subfertiliteit te kunnen onderzoeken. Er werd geen verschil in de NOS, het vloeiend verlopen van bewegingen en het voorkomen van lichte neurologische disfunctie gevonden tussen eenlingen ontstaan na COH-IVF,

MNC-IVF en eenlingen geboren na een natuurlijke conceptie. Opvallend genoeg scoorden de kinderen van de subfertiele paren beter op deze uitkomstmaten dan de referentiegroep. Concluderend is de neurologische conditie van tweejarigen geboren na voortplantingtechnieken en een natuurlijke conceptie vergelijkbaar en lijkt subfertiliteit niet geassocieerd met een slechtere neurologische uitkomst.

In **hoofdstuk 6** worden de mentale en psychomotorische ontwikkeling en het gedrag van de kinderen op de leeftijd van 2 jaar bestudeerd met behulp van de Bayley Ontwikkelingsschalen en Achenbach's kindergedragsvragenlijst. Het effect van subfertiliteit wordt bestudeerd door vergelijking van het cohort met de referentiegroep uit hoofdstuk 5. Er werden geen verschillen gevonden tussen de COH-IVF, MNC-IVF en NC groep. Gedrag en ontwikkeling in de subfertiele groepen was grotendeels vergelijkbaar met de referentiegroep. Als uitzondering scoorden kinderen van subfertiele paren hoger op de schaal betreffende angstig of depressief gedrag. Samenvattend vinden we geen effect van ovariële hyperstimulatie, de IVF procedure of subfertiliteit zelf op cognitie, psychomotore ontwikkeling of gedrag tot op twee jarige leeftijd.

De resultaten van het onderzoek naar de ontwikkeling van de kinderen uit het Groningse cohort zijn geruststellend. Het is echter goed te beseffen dat subtiele neurologische disfunctie op oudere leeftijd kan ontstaan, daarom is het noodzakelijk om het follow-up onderzoek te continueren.

Deel III; Follow-up van kinderen geboren na IVF met Pre-implantatie Genetische Screening

In dit gedeelte van het proefschrift wordt het effect van pre-implantatie genetische screening (PGS) op de neurologische ontwikkeling van het kind onderzocht. Het onderzoek werd uitgevoerd in een samenwerkingsverband tussen het Universitair Medisch Centrum Groningen en het Academisch Medisch Centrum te Amsterdam.

In **hoofdstuk 7** beschrijven we een prospectief follow-up onderzoek van kinderen van paren die deelnamen aan het 'Groningse' deel van het gerandomiseerde onderzoek naar het effect van PGS. De primaire uitkomstmaat van het follow-up onderzoek is een afwijkende neurologische ontwikkeling op de leeftijd van 18 maanden. Secundair zijn neurologische uitkomst voor 18 maanden, het type lichte neurologische disfunctie, opname op de neonatale intensive care (NICU) en prevalentie van aangeboren afwijkingen. Twintig vrouwen in de PGS groep en 26 vrouwen in de controle groep namen deel aan het onderzoek met respectievelijk 25 en 31 kinderen. Vijf zwangerschappen na PGS en 4 zwangerschappen in de controle groep resulteerden in de geboorte van ten minste 1 kind met een afwijkende neurologische uitkomst (25 vs. 15%). Lichte disfunctie in de fijne motoriek en in de tonusregulatie leken vaker voor te komen na PGS. Neurologische

uitkomst voor 18 maanden en aantallen NICU opnames en aangeboren afwijkingen per zwangerschap waren vergelijkbaar in de groepen. Op kindniveau was het percentage ongunstige neurologische uitkomsten echter enigszins hoger na PGS. Alles samen genomen zijn de ontwikkelingsuitkomsten van kinderen geboren na IVF met en zonder PGS in dit onderzoek niet erg verschillend. Als gevolg van de relatief kleine onderzoeksgroep kan echter ook niet geconcludeerd worden dat er geen relatie is tussen PGS en ongunstige ontwikkelingsuitkomsten.

Hoofdstuk 8 beschrijft de resultaten van het follow-up onderzoek van de 'Amsterdamse' en 'Groningse' kinderen geboren na IVF met of zonder PGS op de leeftijd van 2 jaar. Mentale, psychomotorische, neurologische en gedragsmatige uitkomsten van de kinderen zijn geëvalueerd met behulp van de Bayley Ontwikkelingsschalen, het neurologisch onderzoek volgens Hempel en Achenbach's kindergedragsvragenlijst. Vierenvijftig kinderen na IVF met PGS en 77 kinderen na IVF zonder PGS namen deel aan het onderzoek. Er werden geen verschillen gevonden in mentale en psychomotorische ontwikkeling en gedrag tussen de groepen, maar kinderen geboren na IVF met PGS hadden lagere neurologische optimaliteitsscores op de leeftijd van twee. Mediane scores op alle ontwikkelingstesten lagen in de normale range. Concluderend lijkt conceptie middels PGS niet geassocieerd aan ernstige mentale, neurologische en gedragsmatige afwijkingen op twee jarige leeftijd. De lagere neurologische optimaliteitsscores van kinderen geboren na IVF met PGS geven echter aan dat de neurologische ontwikkeling na PGS op de leeftijd van 2 jaar net iets minder gunstig is. Dit zou een voorbode kunnen zijn van een minder gunstige ontwikkelingsneurologische conditie op lange termijn. Het is daarom van belang dat follow-up onderzoek gecontinueerd wordt.

De uitkomsten van de follow-up onderzoeken van kinderen geboren na IVF met PGS benadrukken het belang van de evaluatie van de veiligheid van nieuwe voortplantingstechnieken voordat grootschalige implementatie plaatsvindt.

Deel IV: Algemene discussie en ideeën voor toekomstig onderzoek

In **hoofdstuk 9** worden de bevindingen van het onderzoek in dit proefschrift bediscussieerd. In het Groningse cohort van kinderen geboren na voortplantingstechnieken werden geen nadelige effecten van ovariële hyperstimulatie of de IVF procedure op de neurologische ontwikkeling gevonden tot tweejarige leeftijd. Follow-up van kinderen geboren na IVF met PGS toonde geen relatie tussen conceptie middels PGS en stoornissen in de mentale, psychomotorische en gedragsmatige uitkomst. Er werden echter wel lagere neurologische optimaliteitsscores gevonden bij kinderen in de PGS groep.

Sterke punten van het Groningse cohort onderzoek zijn de unieke

studiegroepen, het lage percentage uitval tijdens de follow-up, de representatieve groep en het gebruik van gedetailleerde, gestandaardiseerde, leeftijdsspecifieke onderzoeksinstrumenten. Het onderzoek wordt beperkt door de relatief kleine onderzoeksgroep en jonge leeftijd van follow-up. Sterke punten van het PGS onderzoek zijn de randomisatie van de subfertiele paren en het hoge percentage deelname aan het follow-up onderzoek. Het belangrijkste minpunt van de studie is de relatief kleine groepen.

Subfertiliteit is geassocieerd met ongunstige perinatale uitkomsten. In ons onderzoek vonden we echter geen consistente relatie tussen neurologische ontwikkeling en subfertiliteit, mogelijk als gevolg van selectiebias. Opheldering van deze relatie zou kunnen bijdragen aan het optimaliseren van de zorg voor zwangeren en zuigelingen.

Toekomstig onderzoek moet zich richten de ontwikkeling van oudere kinderen, adolescenten en tweelingen geboren na voortplantingstechnieken. Daarnaast verdient onderzoek naar cardiovasculaire uitkomsten van kinderen geboren na voortplantingstechnieken aandacht. De veiligheid van nieuwe reproductieve technieken moet geëvalueerd worden voordat grootschalige implementatie van deze technieken plaatsvindt.

References

R

Achenbach TM, Rescorla LA. Manual for ASEBA Preschool Forms & Profiles. Burlington: ASEBA, 2000.
Adamson GD, de Mouzon J, Lancaster P *et al.* World collaborative report on in vitro fertilisation, 2000. *Fertil Steril* 2006; **85**: 1586-622.
Agarwal P, Loh SK, Lim SB *et al.* Two-year neurodevelopmental outcome in children conceived by intracytoplasmic sperm injection: prospective cohort study. *BJOG* 2005; **112**: 1376-83.
Audiens H, Denayer E, Ahlenajafie N *et al.* Sleep apnoea in babies born after in vitro fertilisation. *Eur J Pediatr* 1995; **154**: 454-7.
Australian in vitro fertilisation collaborative group. High incidence of preterm births and early losses in pregnancy after in vitro fertilisation. *Br Med J (Clin Res Ed)* 1985; **291**: 1160-3.
Banerjee I, Shevlin M, Taranissi M *et al.* Health of children conceived after preimplantation genetic diagnosis: a preliminary outcome study. *Reprod Biomed Online* 2008; **16**: 376-81.
Barker DJ. Fetal origins of coronary heart disease. *BMJ* 1995; **311**: 171-4.
Barnes J, Sutcliffe AG, Kristoffersen I *et al.* The influence of assisted reproduction on family functioning and children's socio-emotional development: results from a European study. *Hum Reprod* 2004; **19**: 1480-7.
Batstra L, Neeleman J, Hadders-Algra M. The neurology of learning and behavioural problems in pre-adolescent children. *Acta Psychiatr Scand* 2003; **108**: 92-100.
Bax M, Goldstein M, Rosenbaum P *et al.* Proposed definition and classification of cerebral palsy, April 2005. *Dev Med Child Neurol* 2005; **47**: 571-6.
Bax M, Tydeman C, Flodmark O. Clinical and MRI correlates of cerebral palsy: the European Cerebral Palsy Study. *JAMA* 2006; **296**: 1602-8.
Bayley N. Bayley Scales of Infant Development. 2 edn. San Antonio, USA: The Pychological Corporation, 1993.
Belva F, Henriet S, Liebaers I *et al.* Medical outcome of 8-year-old singleton ICSI children (born >or=32 weeks' gestation) and a spontaneously conceived comparison group. *Hum Reprod* 2007; **22**: 506-15.
Bhutta AT, Cleves MA, Casey PH, Cradock MM, Anand KJ. Cognitive and behavioural outcomes of school-aged children who were born preterm: a meta-analysis. *JAMA* 2002; **288**: 728-37.
Bonduelle M, Liebaers I, Deketelaere V *et al.* Neonatal data on a cohort of 2889 infants born after ICSI (1991-1999) and of 2995 infants born after IVF (1983-1999). *Hum Reprod* 2002; **17** : 671-94.
Bonduelle M, Ponjaert I, Steirteghem AV *et al.* Developmental outcome at 2 years of age for children born after ICSI compared with children born after IVF. *Hum Reprod* 2003; **18**: 342-50.
Bonduelle M, Bergh C, Niklasson A, Palermo GD, Wennerholm UB. Medical follow-up study of 5-year-old ICSI children. *Reprod Biomed Online* 2004; **9**: 91-101.
Bonduelle M, Wennerholm UB, Loft A *et al.* A multi-centre cohort study of the physical health of 5-year-old children conceived after intracytoplasmic sperm injection, in vitro fertilisation and natural conception. *Hum Reprod* 2005; **20**: 413-9.
Bouwstra H, Dijck-Brouwer DA, Wildeman JA *et al.* Long-chain polyunsaturated fatty acids have a positive effect on the quality of general movements of healthy term infants. *Am J Clin Nutr* 2003; **78**: 313-8.
Bouwstra H, Boersma ER, Boehm G *et al.* Exclusive breastfeeding of healthy term infants for at least 6 weeks improves neurological condition. *J Nutr* 2003; **133**: 4243-5.
Bouwstra H, Dijck-Brouwer J, Decsi T *et al.* Neurologic condition of healthy term infants at 18 months: positive association with venous umbilical DHA status and negative association with umbilical trans-fatty acids. *Pediatr Res* 2006; **60**: 334-9.
Bouwstra H, Dijk-Stigter GR, Grooten HMJ *et al.* Prevalence of abnormal general movements in three-months-old infants. *Early Hum Dev* 2009; **85**: 399-403.
Bowen JR, Gibson FL, Leslie GI, Saunders DM. Medical and developmental outcome at 1 year for children conceived by intracytoplasmic sperm injection. *Lancet* 1998; **351**: 1529-34.
Brandes JM, Scher A, Itzkovits J *et al.* Growth and development of children conceived by in vitro fertilisation. *Pediatrics* 1992; **90**: 424-9.
Buck Louis GM, Schisterman EF, Dukic VM, Schieve LA. Research hurdles complicating the analysis of infertility treatment and child health. *Hum Reprod* 2005; **20**: 12-8.
Buitendijk SE. Children after in vitro fertilisation. An overview of the literature. *Int J Technol Assess Health Care* 1999; **15**: 52-65.

Carson C, Kurinczuk JJ, Sacker A *et al.* Cognitive development following ART: effect of choice of comparison group, confounding and mediating factors. *Hum Reprod* 2010; **25**: 244-52.
Ceelen M, Vermeiden JP. Health of human and livestock conceived by assisted reproduction. *Twin Res* 2001; **4**: 412-6.
Ceelen M, van Weissenbruch MM, Vermeiden JP, van Leeuwen FE, Delemarre-van de Waal HA. Cardiometabolic differences in children born after in vitro fertilisation: follow-up study. *J Clin Endocrinol Metab* 2008; **93**: 1682-8.
Colpin H, Demyttenaere K, Vandemeulebroecke L. New reproductive technology and the family: the parent-child relationship following in vitro fertilisation. *J Child Psychol Psychiatry* 1995; **36**: 1429-41.
Colpin H, Soenen S. Parenting and psychosocial development of IVF children: a follow-up study. *Hum Reprod* 2002; **17**: 1116-23.
Colpin H. Parenting and psychological development of IVF children: review of the research literature. *Developmental Review* 2002; **22**: 644-73.
D'Souza SW, Rivlin E, Cadman J *et al.* Children conceived by in vitro fertilisation after fresh embryo transfer. *Arch Dis Child Fetal Neonatal Ed* 1997; **76**: F70-F74.
Dahlquist G, Finnström O, Köster M, Strömberg B. Neurological sequelae in in-vitro fertilisation babies. *Lancet* 2002; **360**: 719.
De Geyter C, De Geyter M, Steimann S, Zhang H, Holzgreve W. Comparative birth weights of singletons born after assisted reproduction and natural conception in previously infertile women. *Hum Reprod* 2006; **21**: 705-12.
De Graaf-Peters VB, Hadders-Algra M. Ontogeny of the human central nervous system: what is happening when? *Early Hum Dev* 2006; **82**: 257-66.
De Mouzon J, Goossens V, Bhattacharya S *et al.* Assisted reproductive technology in Europe, 2006: results generated from European registers by ESHRE. *Hum Reprod* 2010; **25**: 1851-62.
Debrock S, Melotte C, Spiessens C *et al.* Preimplantation genetic screening for aneuploidy of embryos after in vitro fertilisation in women aged at least 35 years: a prospective randomised trial. *Fertil Steril* 2010; **93**: 364-73.
Desmyttere S, De Schepper J, Nekkebroeck J *et al.* Two-year auxological and medical outcome of singletons born after embryo biopsy applied in preimplantation genetic diagnosis or preimplantation genetic screening. *Hum Reprod* 2009; **24**: 470-6.
Dickey RP, Taylor SN, Lu PY *et al.* Spontaneous reduction of multiple pregnancy: incidence and effect on outcome. *Am J Obstet Gynecol* 2002; **186**: 77-83.
Draper ES, Kurinczuk JJ, Abrams KR, Clarke M. Assessment of separate contributions to perinatal mortality of infertility history and treatment: a case-control analysis. *Lancet* 1999; **353**: 1746-9.
Dumoulin JC, Land JA, Van Montfoort AP *et al.* Effect of in vitro culture of human embryos on birthweight of newborns. *Hum Reprod* 2010; **25**: 605-12.
Ericson A, Nygren KG, Olausson PO, Kallen B. Hospital care utilization of infants born after IVF. *Hum Reprod* 2002; **17**: 929-32.
Ertzeid G, Storeng R. The impact of ovarian stimulation on implantation and fetal development in mice. *Hum Reprod* 2001; **16**: 221-5.
Fisher JR, Hammarberg K, Baker GH. Antenatal mood and fetal attachment after assisted conception. *Fertil Steril* 2008; **89**: 1103-12.
Fleuren KM, Smit LS, Stijnen T, Hartman A. New reference values for the Alberta Infant Motor Scale need to be established. *Acta Paediatr* 2007; **96**: 424-7.
Geraedts J, Collins J, Gianaroli L *et al.* What next for preimplantation genetic screening? A polar body approach! *Hum Reprod* 2010; **25**: 575-7.
Gershoni-Baruch R, Scher A, Itskovitz J, Thaler I, Brandes JM. The physical and psychomotor development of children conceived by IVF and exposed to high-frequency vaginal ultrasonography (6.5 MHz) in the first trimester of pregnancy. *Ultrasound Obstet Gynecol* 1991; **1**: 21-8.
Gibson F, McMahon C. Parenting and child psychological development after infertility treatment. *Int Congr Ser* 2004; **1266**: 351-8.

Gibson FL, Ungerer JA, Leslie GI, Saunders DM, Tennant CC. Development, behaviour and temperament: a prospective study of infants conceived through in-vitro fertilisation. *Hum Reprod* 1998; **13**: 1727-32.

Gibson FL, Ungerer JA, McMahon CA, Leslie GI, Saunders DM. The mother-child relationship following in vitro fertilisation (IVF): infant attachment, responsivity, and maternal sensitivity. *J Child Psychol Psychiatry* 2000; **41**: 1015-23.

Gillberg IC, Gillberg C. Children with preschool minor neurodevelopmental disorders. IV: Behaviour and school achievement at age 13. *Dev Med Child Neurol* 1989; **31**: 3-13.

Golombok S, Cook R, Bish A, Murray C. Families created by the new reproductive technologies: quality of parenting and social and emotional development of the children. *Child Dev* 1995; **66**: 285-98.

Golombok S, Brewaeys A, Cook R *et al.* The European study of assisted reproduction families: family functioning and child development. *Hum Reprod* 1996; **11**: 2324-31.

Golombok S, MacCallum F, Goodman E. The "test-tube" generation: parent-child relationships and the psychological well-being of in vitro fertilisation children at adolescence. *Child Dev* 2001; **72**: 599-608.

Golombok S, Brewaeys A, Giavazzi MT *et al.* The European study of assisted reproduction families: the transition to adolescence. *Hum Reprod* 2002; **17**: 830-40.

Golombok S, MacCallum F. Practitioner review: outcomes for parents and children following non-traditional conception: what do clinicians need to know? *J Child Psychol Psychiatry* 2003; **44**: 303-15.

Goossens V, Harton G, Moutou C *et al.* ESHRE PGD Consortium data collection VIII: cycles from January to December 2005 with pregnancy follow-up to October 2006. *Hum Reprod* 2008; **23**: 2629-45.

Griesinger G, Kolibianakis EM, Diedrich K, Ludwig M. Ovarian stimulation for IVF has no quantitative association with birthweight: a registry study. *Hum Reprod* 2008; **23**: 2549-54.

Griffiths R. The Griffiths Mental Development Scales 1996 Revision (Revised by Huntley M). Henley: Association for research in infant and child development, Test agency., 1996.

Groen SE, de Blécourt AC, Postema K, Hadders-Algra M. General movements in early infancy predict neuromotor development at 9 to 12 years of age. *Dev Med Child Neurol* 2005; **47**: 731-8.

Hadders-Algra M, Huisjes HJ, Touwen BC. Perinatal correlates of major and minor neurological dysfunction at school age: a multivariate analysis. *Dev Med Child Neurol* 1988; **30**: 472-81.

Hadders-Algra M, Huisjes HJ, Touwen BC. Perinatal risk factors and minor neurological dysfunction: significance for behaviour and school achievement at nine years. *Dev Med Child Neurol* 1988; **30**: 482-91.

Hadders-Algra M, Groothuis AM. Quality of general movements in infancy is related to neurological dysfunction, ADHD, and aggressive behaviour. *Dev Med Child Neurol* 1999; **41**: 381-91.

Hadders-Algra M. Two distinct forms of minor neurological dysfunction: perspectives emerging from a review of data of the Groningen Perinatal Project. *Dev Med Child Neurol* 2002; **44**: 561-71.

Hadders-Algra M. Developmental coordination disorder: is clumsy motor behaviour caused by a lesion of the brain at early age? *Neural Plast* 2003; **10**: 39-50.

Hadders-Algra M, Mavinkurve-Groothuis AM, Groen SE *et al.* Quality of general movements and the development of minor neurological dysfunction at toddler and school age. *Clin Rehabil* 2004; **18**: 287-99.

Hadders-Algra M. General movements: A window for early identification of children at high risk for developmental disorders. *J Pediatr* 2004; **145**: S12-S18.

Hadders-Algra M. The neuromotor examination of the preschool child and its prognostic significance. *Ment Retard Dev Disabil Res Rev* 2005; **11**: 180-8.

Hadders-Algra M. Atypical performance: how do we deal with that? *Dev Med Child Neurol* 2007; **49**: 403.

Hadders-Algra M, Bouwstra H, Groen SE. Quality of general movements and psychiatric morbidity at 9 to 12 years. *Early Hum Dev* 2008.

Hadders-Algra M, Heineman KR, Bos AF, Middelburg KJ. The assessment of minor neurological dysfunction in infancy using the Touwen Infant Neurological Examination: strengths and limitations. *Dev Med Child Neurol* 2009.

Hadders-Algra M. The examination of the child with minor neurological dysfunction. 3rd edn. Mac Keith Press, 2010.
Hahn CS. Review: psychosocial well-being of parents and their children born after assisted reproduction. *J Pediatr Psychol* 2001; **26**: 525-38.
Hahn CS, DiPietro JA. In vitro fertilisation and the family: quality of parenting, family functioning, and child psychosocial adjustment. *Dev Psychol* 2001; **37**: 37-48.
Hansen M, Colvin L, Petterson B *et al.* Admission to hospital of singleton children born following assisted reproductive technology (ART). *Hum Reprod* 2008; **23**: 1297-305.
Hardarson T, Hanson C, Lundin K *et al.* Preimplantation genetic screening in women of advanced maternal age caused a decrease in clinical pregnancy rate: a randomised controlled trial. *Hum Reprod* 2008; **23**: 2806-12.
Harper J, Coonen E, De Rycke M *et al.* What next for preimplantation genetic screening (PGS)? A position statement from the ESHRE PGD Consortium steering committee. *Hum Reprod* 2010.
Heineman KR, Hadders-Algra M. Evaluation of neuromotor function in infancy-A systematic review of available methods. *J Dev Behav Pediatr* 2008; **29**: 315-23.
Helmerhorst FM, Perquin DA, Donker D, Keirse MJ. Perinatal outcome of singletons and twins after assisted conception: a systematic review of controlled studies. *BMJ* 2004; **328**: 261-4.
Hempel MS. The neurological examination for toddler-age. Groningen, The Netherlands: University of Groningen, 1993.
Huisman M, Koopman-Esseboom C, Lanting CI *et al.* Neurological condition in 18-month-old children perinatally exposed to polychlorinated biphenyls and dioxins. *Early Hum Dev* 1995; **43**: 165-76.
Hvidtjørn D, Grove J, Schendel DE *et al.* Cerebral palsy among children born after in vitro fertilisation: the role of preterm delivery--a population-based, cohort study. *Pediatrics* 2006; **118**: 475-82.
Hvidtjørn D, Schieve L, Schendel D *et al.* Cerebral Palsy, Autism Spectrum Disorders, and Developmental Delay in Children Born After Assisted Conception: A Systematic Review and Meta-analysis. *Arch Pediatr Adolesc Med* 2009; **163**: 72-83.
Jackson RA, Gibson KA, Wu YW, Croughan MS. Perinatal outcomes in singletons following in vitro fertilisation: a meta-analysis. *Obstet Gynecol* 2004; **103**: 551-63.
Jansen RP, Bowman MC, de Boer KA *et al.* What next for preimplantation genetic screening (PGS)? Experience with blastocyst biopsy and testing for aneuploidy. *Hum Reprod* 2008; **23**: 1476-8.
Kanyo K, Konc J. A follow-up study of children born after diode laser assisted hatching. *Eur J Obstet Gynecol Reprod Biol* 2003; **110**: 176-80.
Kapiteijn K, de Bruijn CS, de Boer E *et al.* Does subfertility explain the risk of poor perinatal outcome after IVF and ovarian hyperstimulation? *Hum Reprod* 2006; **21**: 3228-34.
Källén B, Finnstrom O, Nygren KG, Olausson PO. In vitro fertilisation (IVF) in Sweden: infant outcome after different IVF fertilisation methods. *Fertil Steril* 2005; **84**: 611-7.
Källén B, Finnström O, Nygren KG, Olausson PO. In vitro fertilisation in Sweden: child morbidity including cancer risk. *Fertil Steril* 2005; **84**: 605-10.
Khosla S, Dean W, Brown D, Reik W, Feil R. Culture of preimplantation mouse embryos affects fetal development and the expression of imprinted genes. *Biol Reprod* 2001; **64**: 918-26.
Klemetti R, Sevon T, Gissler M, Hemminki E. Health of children born as a result of in vitro fertilisation. *Pediatrics* 2006; **118**: 1819-27.
Klemetti R, Sevon T, Gissler M, Hemminki E. Health of children born after ovulation induction. *Fertil Steril* 2010; **93**: 1157-68.
Knoester M. Development and Health after ICSI, Development and Health of 5-8-year-old Singletons born after Intracytoplsmic Sperm Injection. Leiden, The Netherlands: Leiden University Medical Center, 2007.
Knoester M, Vandenbroucke JP, Helmerhorst FM *et al.* Matched follow-up study of 5-8 year old ICSI-singletons: comparison of their neuromotor development to IVF and naturally conceived singletons. *Hum Reprod* 2007; **22**: 1638-46.
Knoester M, Helmerhorst FM, van der Westerlaken LA, Walther FJ, Veen S. Matched follow-up study of 5 8-year-old ICSI singletons: child behaviour, parenting stress and child (health-related) quality of life. *Hum Reprod* 2007; **22**: 3098-107.

Knoester M, Helmerhorst FM, Vandenbroucke JP *et al.* Cognitive development of singletons born after intracytoplasmic sperm injection compared with in vitro fertilisation and natural conception. *Fertil Steril* 2008; **90**: 289-96.

Koivurova S, Hartikainen AL, Sovio U *et al.* Growth, psychomotor development and morbidity up to 3 years of age in children born after IVF. *Hum Reprod* 2003; **18**: 2328-36.

La Sala GB, Gallinelli A, Fagandini P *et al.* Developmental outcomes at one and two years of children conceived by intracytoplasmic sperm injection. *Int J Fertil Womens Med* 2004; **49**: 113-9.

Lambert RD. Safety issues in assisted reproductive technology: aetiology of health problems in singleton ART babies. *Hum Reprod* 2003; **18**: 1987-91.

Land JA, Evers JL. What is the most relevant standard of success in assisted reproduction? Defining outcome in ART: a Gordian knot of safety, efficacy and quality. *Hum Reprod* 2004; **19** : 1046-8.

Largo RH, Pfister D, Molinari L *et al.* Significance of prenatal, perinatal and postnatal factors in the development of AGA preterm infants at five to seven years. *Dev Med Child Neurol* 1989; **31**: 440-56.

Leslie GI, Gibson FL, McMahon C *et al.* Children conceived using ICSI do not have an increased risk of delayed mental development at 5 years of age. *Hum Reprod* 2003; **18**: 2067-72.

Leslie GI. Mental development of children conceived using intracytoplasmic sperm injection. The current evidence. *Minerva Ginecol* 2004; **56**: 247-57.

Leunens L, Celestin-Westreich S, Bonduelle M, Liebaers I, Ponjaert-Kristoffersen I. Cognitive and motor development of 8-year-old children born after ICSI compared to spontaneously conceived children. *Hum Reprod* 2006; **21**: 2922-9.

Leunens L, Celestin-Westreich S, Bonduelle M, Liebaers I, Ponjaert-Kristoffersen I. Follow-up of cognitive and motor development of 10-year-old singleton children born after ICSI compared with spontaneously conceived children. *Hum Reprod* 2008; **23**: 105-11.

Levy-Shiff R, Vakil E, Dimitrovsky L *et al.* Medical, cognitive, emotional, and behavioural outcomes in school-age children conceived by in-vitro fertilisation. *J Clin Child Psychol* 1998; **27**: 320-9.

Lidegaard O, Pinborg A, Andersen AN. Imprinting diseases and IVF: Danish National IVF cohort study. *Hum Reprod* 2005; **20**: 950-4.

Ludwig A, Katalinic A, Thyen U *et al.* Neuromotor development and mental health at 5.5 years of age of singletons born at term after intracytoplasmatic sperm injection ICSI: results of a prospective controlled single-blinded study in Germany. *Fertil Steril* 2009; **91**: 125-32.

Ludwig AK, Sutcliffe AG, Diedrich K, Ludwig M. Post-neonatal health and development of children born after assisted reproduction: a systematic review of controlled studies. *Eur J Obstet Gynecol Reprod Biol* 2006; **127**: 3-25.

Ludwig AK, Katalinic A, Entenmann A *et al.* Can we sense ART? The blinded examiner is not blind-a problem with follow-up studies on children born after assisted reproduction. *Fertil Steril* 2009.

Ludwig M, Diedrich K. Follow-up of children born after assisted reproductive technologies. *Reprod Biomed Online* 2002; **5**: 317-22.

Maimburg RD, Vaeth M. Do children born after assisted conception have less risk of developing infantile autism? *Hum Reprod* 2007; **22**: 1841-3.

Mastenbroek S, Twisk M, Echten-Arends J *et al.* In vitro fertilisation with preimplantation genetic screening. *N Engl J Med* 2007; **357**: 9-17.

Mastenbroek S, Scriven P, Twisk M *et al.* What next for preimplantation genetic screening? More randomised controlled trials needed? *Hum Reprod* 2008; **23**: 2626-8.

McMahon C, Gibson F. A special path to parenthood: parent-child relationships in families giving birth to singleton infants through IVF. *Reprod Biomed Online* 2002; **5**: 179-86.

McMahon CA, Ungerer JA, Tennant C, Saunders D. Psychosocial adjustment and the quality of the mother-child relationship at four months postpartum after conception by in vitro fertilisation. *Fertil Steril* 1997; **68**: 492-500.

Meyer LR, Klipstein S, Hazlett WD *et al.* A prospective randomised controlled trial of preimplantation genetic screening in the "good prognosis" patient. *Fertil Steril* 2009; **91**: 1731-8.

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. *Hum Reprod Update* 2008; **14**: 219-31.

Middelburg KJ, Heineman MJ, Bos AF *et al.* The Groningen ART cohort study: ovarian hyperstimulation and the in vitro procedure do not affect neurological outcome in infancy. *Hum Reprod* 2009; **24**: 3119-26.

Middelburg KJ, Haadsma ML, Heineman MJ, Bos AF, Hadders-Algra M. Ovarian hyperstimulation and the in vitro fertilisation procedure do not influence early neuromotor development; a history of subfertility does. *Fertil Steril* 2010; **93**: 544-53.

Middelburg KJ, Heineman MJ, Haadsma ML *et al.* Neurological condition of infants born after in vitro fertilisation with preimplantation genetic screening. *Pediatr Res* 2010.

Moher D, Cook DJ, Eastwood S *et al.* Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet* 1999; **354**: 1896-900.

Morin NC, Wirth FH, Johnson DH *et al.* Congenital malformations and psychosocial development in children conceived by in vitro fertilisation. *J Pediatr* 1989; **115**: 222-7.

Moster D, Lie RT, Markestad T. Long-term medical and social consequences of preterm birth. *N Engl J Med* 2008; **359**: 262-73.

Munne S, Weier HU, Stein J, Grifo J, Cohen J. A fast and efficient method for simultaneous X and Y in situ hybridization of human blastomeres. *J Assist Reprod Genet* 1993; **10**: 82-90.

Munne S, Lee A, Rosenwaks Z, Grifo J, Cohen J. Diagnosis of major chromosome aneuploidies in human preimplantation embryos. *Hum Reprod* 1993; **8**: 2185-91.

Nargund G, Fauser BC, Macklon NS *et al.* The ISMAAR proposal on terminology for ovarian stimulation for IVF. *Hum Reprod* 2007; **22**: 2801-4.

Nekkebroeck J, Bonduelle M, Desmyttere S, Van den Broeck W, Ponjaert-Kristoffersen I. Socio-emotional and language development of 2-year-old children born after PGD/PGS, and parental well-being. *Hum Reprod* 2008; **23**: 1849-57.

Nekkebroeck J, Bonduelle M, Desmyttere S, Van den Broeck W, Ponjaert-Kristoffersen I. Mental and psychomotor development of 2-year-old children born after preimplantation genetic diagnosis/screening. *Hum Reprod* 2008; **23**: 1560-6.

Neri QV, Tanaka N, Wang A *et al.* Intracytoplasmic sperm injection. Accomplishments and qualms. *Minerva Ginecol* 2004; **56**: 189-96.

Nyboe Andersen A, Goossens V, Gianaroli L *et al.* Assisted reproductive technology in Europe, 2003. Results generated from European registers by ESHRE. *Hum Reprod* 2007; **22**: 1513-25.

Nyboe Andersen A, Goossens V, Ferraretti AP *et al.* Assisted reproductive technology in Europe, 2004: results generated from European registers by ESHRE. *Hum Reprod* 2008; **23**: 756-71.

Nyboe Andersen A, Goossens V, Bhattacharya S *et al.* Assisted reproductive technology and intrauterine inseminations in Europe, 2005: results generated from European registers by ESHRE: ESHRE. The European IVF Monitoring Programme (EIM), for the European Society of Human Reproduction and Embryology (ESHRE). *Hum Reprod* 2009; **24**: 1267-87.

Olivennes F, Rufat P, André B *et al.* The increased risk of complication observed in singleton pregnancies resulting from in-vitro fertilisation (IVF) does not seem to be related to the IVF method itself. *Hum Reprod* 1993; **8**: 1297-300.

Olivennes F, Fanchin R, Ledee N *et al.* Perinatal outcome and developmental studies on children born after IVF. *Hum Reprod Update* 2002; **8**: 117-28.

Olivennes F, Golombok S, Ramogida C, Rust J. Behavioural and cognitive development as well as family functioning of twins conceived by assisted reproduction: findings from a large population study. *Fertil Steril* 2005; **84**: 725-33.

Ombelet W, Cadron I, Gerris J *et al.* Obstetric and perinatal outcome of 1655 ICSI and 3974 IVF singleton and 1102 ICSI and 2901 IVF twin births: a comparative analysis. *Reprod Biomed Online* 2005; **11**: 76-85.

Ombelet W, Martens G, De Sutter P *et al.* Perinatal outcome of 12,021 singleton and 3108 twin births after non-IVF-assisted reproduction: a cohort study. *Hum Reprod* 2006; **21**: 1025-32.

Painter RC, Roseboom TJ. Cardiovascular health among children born after assisted reproduction. *Eur J Obstet Gynecol Reprod Biol* 2007; **131**: 107-8.

Palermo G, Joris H, Devroey P, Van Steirteghem AC. Pregnancies after intracytoplasmic injection of single spermatozoon into an oocyte. *Lancet* 1992; **340**: 17-8.

Pandian Z, Bhattacharya S, Templeton A. Review of unexplained infertility and obstetric outcome: a 10 year review. *Hum Reprod* 2001; **16**: 2593-7.

Papaligoura Z, Panopoulou-Maratou O, Solman M, Arvaniti K, Sarafidou J. Cognitive development of 12 month old Greek infants conceived after ICSI and the effects of the method on their parents. *Hum Reprod* 2004; **19**: 1488-93.

Pelinck MJ, Vogel NE, Hoek A *et al.* Minimal stimulation IVF with late follicular phase administration of the GnRH antagonist cetrorelix and concomitant substitution with recombinant FSH: a pilot study. *Hum Reprod* 2005; **20**: 642-8.

Pelinck MJ, Vogel NE, Arts EG *et al.* Cumulative pregnancy rates after a maximum of nine cycles of modified natural cycle IVF and analysis of patient drop-out: a cohort study. *Hum Reprod* 2007; **22**: 2463-70.

Pelinck MJ, Knol HM, Vogel NE *et al.* Cumulative pregnancy rates after sequential treatment with modified natural cycle IVF followed by IVF with controlled ovarian stimulation. *Hum Reprod* 2008; **23**: 1808-14.

Pelinck MJ, Keizer MH, Hoek A *et al.* Perinatal outcome in singletons after modified natural cycle IVF and standard IVF with ovarian stimulation. *Eur J Obstet Gynecol Reprod Biol* 2010; **148**: 56-61.

Pennings G, de Wert G, Shenfield F *et al.* ESHRE Task Force on Ethics and Law 13: the welfare of the child in medically assisted reproduction. *Hum Reprod* 2007; **22**: 2585-8.

Pinborg A, Loft A, Schmidt L, Andersen AN. Morbidity in a Danish national cohort of 472 IVF/ICSI twins, 1132 non-IVF/ICSI twins and 634 IVF/ICSI singletons: health-related and social implications for the children and their families. *Hum Reprod* 2003; **18**: 1234-43.

Pinborg A, Loft A, Schmidt L *et al.* Neurological sequelae in twins born after assisted conception: controlled national cohort study. *BMJ* 2004; **329**: 311.

Pinborg A, Lidegaard O, la Cour Freiesleben N, Andersen AN. Consequences of vanishing twins in IVF/ICSI pregnancies. *Hum Reprod* 2005; **20**: 2821-9.

Pinborg A. IVF/ICSI twin pregnancies: risks and prevention. *Hum Reprod Update* 2005; **11**: 575-93.

Pinborg A, Lidegaard O, la Cour Freiesleben N, Andersen AN. Vanishing twins: a predictor of small-for-gestational age in IVF singletons. *Hum Reprod* 2007; **22**: 2707-14.

Pinborg A, Loft A, Aaris Henningsen AK, Rasmussen S, Nyboe Andersen A. Infant outcome of 957 singletons born after frozen embryo replacement: The Danish National Cohort Study 1995-2006. *Fertil Steril* 2009.

Place I, Englert Y. A prospective longitudinal study of the physical, psychomotor, and intellectual development of singleton children up to 5 years who were conceived by intracytoplasmic sperm injection compared with children conceived spontaneously and by in vitro fertilisation. *Fertil Steril* 2003; **80**: 1388-97.

Ponjaert-Kristoffersen I, Tjus T, Nekkebroeck J *et al.* Psychological follow-up study of 5-year-old ICSI children. *Hum Reprod* 2004; **19**: 2791-7.

Ponjaert-Kristoffersen I, Bonduelle M, Barnes J *et al.* International collaborative study of intracytoplasmic sperm injection-conceived, in vitro fertilisation-conceived, and naturally conceived 5-year-old child outcomes: cognitive and motor assessments. *Pediatrics* 2005; **115**: e283-e289.

Prechtl HF. The optimality concept. *Early Hum Dev* 1980; **4**: 201-5.

Prechtl HF. Qualitative changes of spontaneous movements in fetus and preterm infant are a marker of neurological dysfunction. *Early Hum Dev* 1990; **23**: 151-8.

Pridham KF, Brown R, Clark R, Sondel S, Green C. Infant and caregiving factors affecting weight-for-age and motor development of full-term and premature infants at 1 year post-term. *Res Nurs Health* 2002; **25**: 394-410.

Raoul-Duval A, Bertrand-Servais M, Frydman R. Comparative prospective study of the psychological development of children born by in vitro fertilisation and their mothers. *J Psychosom Obstet Gynaecol* 1993; **14**: 117-26.

Raoul-Duval A, Bertrand-Servais M, Letur-Konirsch H, Frydman R. Psychological follow-up of children born after in-vitro fertilisation. *Hum Reprod* 1994; **9**: 1097-101.

Repokari L, Punamaki RL, Poikkeus P *et al.* The impact of successful assisted reproduction treatment on female and male mental health during transition to parenthood: a prospective controlled study. *Hum Reprod* 2005; **20**: 3238-47.

Rerat M, Zbinden Y, Saner R, Hammon H, Blum JW. In vitro embryo production: growth performance, feed efficiency, and hematological, metabolic, and endocrine status in calves. *J Dairy Sci* 2005; **88**: 2579-93.

Romundstad LB, Romundstad PR, Sunde A *et al.* Effects of technology or maternal factors on perinatal outcome after assisted fertilisation: a population-based cohort study. *Lancet* 2008; **372**: 737-43.

Ron-El R, Lahat E, Golan A *et al.* Development of children born after ovarian superovulation induced by long-acting gonadotropin-releasing hormone agonist and menotropins, and by in vitro fertilisation. *J Pediatr* 1994; **125**: 734-7.

Rongières-Bertrand C, Olivennes F, Righini C *et al.* Revival of the natural cycles in in-vitro fertilisation with the use of a new gonadotrophin-releasing hormone antagonist (Cetrorelix): a pilot study with minimal stimulation. *Hum Reprod* 1999; **14**: 683-8.

Ruiter SAJ, Spelberg HCL, van der Meulen BF, Nakken H. The BSID-II-NL: construction, standardisation and instrument utility. *Netherlands Journal of Psychology* 2008; **64**: 15-40.

Salem Yaniv S., Levy A, Wiznitzer A *et al.* A significant linear association exists between advanced maternal age and adverse perinatal outcome. *Arch Gynecol Obstet* 2011; **283**: 755-9.

Sanchez-Albisua I, Borell-Kost S, Mau-Holzmann UA, Licht P, Krageloh-Mann I. Increased frequency of severe major anomalies in children conceived by intracytoplasmic sperm injection. *Dev Med Child Neurol* 2007; **49**: 129-34.

Saunders K, Spensley J, Munro J, Halasz G. Growth and physical outcome of children conceived by in vitro fertilisation. *Pediatrics* 1996; **97**: 688-92.

Schieve LA, Meikle SF, Ferre C *et al.* Low and very low birth weight in infants conceived with use of assisted reproductive technology. *N Engl J Med* 2002; **346**: 731-7.

Schoolcraft WB, Katz-Jaffe MG, Stevens J, Rawlins M, Munne S. Preimplantation aneuploidy testing for infertile patients of advanced maternal age: a randomised prospective trial. *Fertil Steril* 2009; **92**: 157-62.

Sermon K, Van Steirteghem A, Liebaers I. Preimplantation genetic diagnosis. *Lancet* 2004; **363**: 1633-41.

Soorani-Lunsing RJ, Hadders-Algra M, Huisjes HJ, Touwen BC. Minor neurological dysfunction after the onset of puberty: association with perinatal events. *Early Hum Dev* 1993; **33**: 71-80.

Staessen C, Platteau P, Van Assche E *et al.* Comparison of blastocyst transfer with or without preimplantation genetic diagnosis for aneuploidy screening in couples with advanced maternal age: a prospective randomised controlled trial. *Hum Reprod* 2004; **19**: 2849-58.

Staessen C, Verpoest W, Donoso P *et al.* Preimplantation genetic screening does not improve delivery rate in women under the age of 36 following single-embryo transfer. *Hum Reprod* 2008; **23**: 2818-25.

Statistics Netherlands Website. StatLine databank. November 3, 2008. http://statline.cbs.nl/StatWeb/default.aspx.

Steptoe PC, Edwards RG. Birth after the reimplantation of a human embryo. *Lancet* 1978; **2**: 366.

Strömberg B, Dahlquist G, Ericson A *et al.* Neurological sequelae in children born after in-vitro fertilisation: a population-based study. *Lancet* 2002; **359**: 461-5.

Sun Y, Vestergaard M, Christensen J *et al.* Epilepsy and febrile seizures in children of treated and untreated subfertile couples. *Hum Reprod* 2007; **22**: 215-20.

Sutcliffe AG, D'Souza SW, Cadman J *et al.* Minor congenital anomalies, major congenital malformations and development in children conceived from cryopreserved embryos. *Hum Reprod* 1995; **10**: 3332-7.

Sutcliffe AG, D'Souza SW, Cadman J *et al.* Outcome in children from cryopreserved embryos. *Arch Dis Child* 1995; **72**: 290-3.

Sutcliffe AG, Taylor B, Grudzinskas G, Thornton S, Lieberman B. Children conceived by intracytoplasmic sperm injection. *Lancet* 1998; **352**: 578-9.

Sutcliffe AG, Taylor B, Li J *et al.* Children born after intracytoplasmic sperm injection: population control study. *BMJ* 1999; **318**: 704-5.

Sutcliffe AG. Follow-up of children conceived from cryopreserved embryos. *Mol Cell Endocrinol* 2000; **169** : 91-3.
Sutcliffe AG, Taylor B, Saunders K *et al.* Outcome in the second year of life after in-vitro fertilisation by intracytoplasmic sperm injection: a UK case-control study. *Lancet* 2001; **357**: 2080-4.
Sutcliffe AG, Saunders K, McLachlan R *et al.* A retrospective case-control study of developmental and other outcomes in a cohort of Australian children conceived by intracytoplasmic sperm injection compared with a similar group in the United Kingdom. *Fertil Steril* 2003; **79**: 512-6.
Sutcliffe AG, Edwards P, Beeson C, Barnes J. Comparing parents' perceptions of IVF conceived children's behaviour with naturally conceived children. *Infant Ment Health J* 2004; **25**: 163-70.
Sutcliffe AG, Barnes J, Wennerholm UB *et al.* Laterality in five-year-olds conceived by intracytoplasmic sperm injection, standard in vitro fertilisation and natural conception: a European study. *BJOG* 2005; **112**: 1397-401.
Sutcliffe AG, Ludwig M. Outcome of assisted reproduction. *Lancet* 2007; **370**: 351-9.
Sydsjo G, Wadsby M, Kjellberg S, Sydsjo A. Relationships and parenthood in couples after assisted reproduction and in spontaneous primiparous couples: a prospective long-term follow-up study. *Hum Reprod* 2002; **17**: 3242-50.
Tarlatzis BC, Grimbizis G. Pregnancy and child outcome after assisted reproduction techniques. *Hum Reprod* 1999; **14**: 231-42.
Te Velde ER, van Baar AL, van Kooij RJ. Concerns about assisted reproduction. *Lancet* 1998; **351**: 1524-5.
Thomson F, Shanbhag S, Templeton A, Bhattacharya S. Obstetric outcome in women with subfertility. *BJOG* 2005; **112**: 632-7.
Touwen BCL. Neurological development in infancy. Philadelphia: Lippincott., 1976.
Treyvaud K, Anderson VA, Howard K *et al.* Parenting behaviour is associated with the early neurobehavioural development of very preterm children. *Pediatrics* 2009; **123**: 555-61.
Ulrich D, Gagel DE, Hemmerling A, Pastor VS, Kentenich H. Couples becoming parents: something special after IVF? *J Psychosom Obstet Gynaecol* 2004; **25**: 99-113.
Van Balen F. Child-rearing following in vitro fertilisation. *J Child Psychol Psychiatry* 1996; **37**: 687-93.
Van Balen F. Development of IVF children. *Dev Rev* 1998; **18**: 30-46.
Van der Auwera I, D'Hooghe T. Superovulation of female mice delays embryonic and fetal development. *Hum Reprod* 2001; **16**: 1237-43.
Van der Meulen BF, Ruiter SAJ, Spelberg HCL, Smrkovsky M. Bayley Scales of Infant Development- II. Dutch version. 2 edn. Lisse, The Netherlands: Smets and Zeitlinger B.V., 2004.
Van Steirteghem A, Bonduelle M, Devroey P, Liebaers I. Follow-up of children born after ICSI. *Hum Reprod Update* 2002; **8**: 111-6.
Van Wely M, Twisk M, Mol BW, Van d, V. Is twin pregnancy necessarily an adverse outcome of assisted reproductive technologies? *Hum Reprod* 2006; **21**: 2736-8.
Wagenaar K, Ceelen M, van Weissenbruch MM *et al.* School functioning in 8- to 18-year-old children born after in vitro fertilisation. *Eur J Pediatr* 2008; **167**: 1289-95.
Wagenaar K, van Weissenbruch MM, Knol DL *et al.* Behaviour and socioemotional functioning in 9-18-year-old children born after in vitro fertilisation. *Fertil Steril* 2009; **92**: 1907-14.
Wang YA, Sullivan EA, Black D *et al.* Preterm birth and low birth weight after assisted reproductive technology-related pregnancy in Australia between 1996 and 2000. *Fertil Steril* 2005; **83**: 1650-8.
Watkins AJ, Platt D, Papenbrock T *et al.* Mouse embryo culture induces changes in postnatal phenotype including raised systolic blood pressure. *Proc Natl Acad Sci U S A* 2007; **104**: 5449-54.
Wennerholm UB, Albertsson-Wikland K, Bergh C *et al.* Postnatal growth and health in children born after cryopreservation as embryos. *Lancet* 1998; **351**: 1085-90.
Wennerholm UB, Bergh C, Hamberger L *et al.* Obstetric outcome of pregnancies following ICSI, classified according to sperm origin and quality. *Hum Reprod* 2000; **15**: 1189-94.
Wennerholm WB. Cryopreservation of embryos and oocytes: obstetric outcome and health in children. *Hum Reprod* 2000; **15 Suppl 5**: 18-25.
Wikstrand MH, Stromland K, Flodin S *et al.* Ophthalmological findings in children born after intracytoplasmic sperm injection. *Acta Ophthalmol Scand* 2006; **84**: 177-81.

Wilton L. Preimplantation genetic diagnosis for aneuploidy screening in early human embryos: a review. *Prenat Diagn* 2002; **22**: 512-8.

Wright VC, Chang J, Jeng G, Macaluso M. Assisted reproductive technology surveillance--United States, 2005. *MMWR Surveill Summ* 2008; **57**: 1-23.

Young LE, Fernandes K, McEvoy TG *et al.* Epigenetic change in IGF2R is associated with fetal overgrowth after sheep embryo culture. *Nat Genet* 2001; **27**: 153-4.

Zeilmaker GH, Alberda AT, van G, I, Rijkmans CM, Drogendijk AC. Two pregnancies following transfer of intact frozen-thawed embryos. *Fertil Steril* 1984; **42**: 293-6.

Zhu JL, Basso O, Obel C, Hvidtjorn D, Olsen J. Infertility, infertility treatment and psychomotor development: the Danish National Birth Cohort. *Paediatr Perinat Epidemiol* 2009; **23**: 98-106.

Zhu JL, Obel C, Basso O, Olsen J. Parental infertility and developmental coordination disorder in children. *Hum Reprod* 2010; **25**: 908-13.

Zhu JL, Hvidtjorn D, Basso O *et al.* Parental infertility and cerebral palsy in children. *Hum Reprod* 2010; **25**: 3142-5.

Dankwoord

Over de auteur

List of publications

DANKWOORD

Dit proefschrift is tot stand gekomen door samenwerking tussen de afdelingen Ontwikkelingsneurologie, Voortplantingsgeneeskunde en Neonatologie van het Universitair Medisch Centrum Groningen. Mijn dank gaat uit naar iedereen die direct of indirect betrokken is geweest bij het onderzoek. In het bijzonder wil ik prof. dr. M. Hadders-Algra, prof. dr. M.J. Heineman en prof. dr. A.F. Bos bedanken.

Beste Mijna, jij hebt me wegwijs gemaakt in de wetenschap en de ontwikkelingsneurologie. Jouw kennis en deskundigheid zijn een enorme inspiratiebron voor mij. Jouw kritische houding heeft altijd tot betere manuscripten geleid en ik ben jou daar dan ook zeer erkentelijk voor. We hebben samen urenlang bewegingspatronen bekeken, bediscussieerd en beoordeeld. Jouw bevlogen analyses en voorbeelden maakten de stof inzichtelijk. De liters thee, koffie en jouw schaterlach maakten die uren ook gezellig. Jouw enorme persoonlijke inzet, oprechtheid, doorzettingsvermogen en supersnelle reacties maakten onze samenwerking voor mij erg plezierig. Ik verheug me dan ook enorm op de manuscripten die nog gaan komen.

Beste Maas Jan, ik weet nog goed dat ik als student bij je langs kwam voor een wetenschappelijke stage. Jij maakte me enthousiast voor de uitkomsten van hulp bij voortplanting, maar ik had nooit verwacht dat dit in een proefschrift zou resulteren! Jouw rust, analyserend vermogen en persoonlijke benadering zijn onmisbaar geweest voor het onderzoek. Jij hielp me relativeren en stond altijd klaar met positieve en bemoedigende woorden als ik soms even dacht dat het onmogelijk was alle kinderen in de daarvoor beschikbare tijd te zien. Ook na jouw vertrek naar Amsterdam ben je zeer betrokken gebleven en nam je geregeld de moeite om vergaderingen in Groningen bij te wonen. Geweldig vind ik dat. Het is opnieuw anders gelopen dan ik had verwacht (nog beter!); nu ben je mijn opleider in het Amsterdamse. Ik kijk er erg naar uit om de komende jaren met jou als opleider mijn opleiding in het AMC te vervolgen.

Beste Arie, jij wist naast je overvolle kliniek altijd tijd vrij te maken om eens rustig te discussiëren over oorzaken, dan wel gevolgen en onderzoeksmethoden. Jij bood tegengewicht, maar veerde mee waar nodig. Veel dank voor je kritisch meeschrijven, je klinische blik op het onderzoek en de tijd die je in het follow-up onderzoek hebt gestoken.

De leden van de beoordelingscommissie; prof. dr. F.M. Helmerhorst, prof. dr. J.H. Kok en prof. dr. C.M. Bilardo wil ik hartelijk danken voor het beoordelen en goedkeuren van dit proefschrift.

Alle ouders en kinderen die hebben deelgenomen aan het onderzoek wil ik graag enorm bedanken voor hun bereidheid, hun vertrouwen en hun toewijding aan de studies die staan beschreven in dit proefschrift. Het was geweldig om onderzoek met jullie te mogen doen.

Omdat kleine kinderen groot worden heeft dit onderzoek een vervolg gekregen op 4 jarige leeftijd. Marjolein Jongbloed-Pereboom kwam als eerste om me op te volgen en te helpen met de soms oneindig lijkende stroom peuters. Wat was ik blij met jouw hulp! (Lees: het separaat entertainen van tweejarige tweelingen!) Je hebt het geweldig gedaan. Super, dat je nu in Nijmegen je eigen promotieonderzoek hebt. Vervolgens kwamen Pamela Schendelaar en Jorien Seggers onze onderzoeksgroep versterken. Ik vind het een topprestatie dat jullie naast je coschappen het onderzoek draaiende weten te houden. Jullie nieuwsgierigheid, enthousiasme, inspanning en doorzettingsvermogen gaat zeker nog tot een hele reeks prachtige publicaties leiden.

Zonder alle studenten die een wetenschappelijke stage bij ons hebben gedaan was dit onderzoek nooit tot een succes geworden. Jolanda Winter, Jan Groenewegen, Wendy Nijhuis, Susanne Kiewiet, Joke Duursma, Jolien Snijders, Hilde Bangma, Martje Dijkstra, Eva Loomans, Mejgan Jelani, Inonge Reimert, Marijke Thunnisen, Aisha Sinot, Jolanda Steenwijk & Madelein van der Stouwe, ik ben jullie heel veel dank verschuldigd voor jullie hulp bij het onderzoeken van de kinderen. Ik zal de soms ietwat gehaaste ritjes in de provincie, gewapend met een camera, nooit vergeten.

Collega's van de afdeling voortplantingsgeneeskunde: dr. A. Hoek, dr. M.J. Pelinck en dr. J. van Echten. Beste Annemieke, Marie José en Jannie, dank dat jullie mij kennis hebben laten maken met de wereld van de wetenschap en in een later stadium voor de gezamenlijke begeleiding van studenten en de steun en gezelligheid bij congresbezoek. Prof. dr. J. Land, bedankt voor het belangeloos kritisch doornemen van manuscripten en het zitting nemen in de corona.

Dr. Fidler, beste Vaclav, dank voor je hulp bij de ingewikkelde statistische analyses. Door jouw geduldige uitleg werden statistiek en methodologie begrijpelijk en steeds leuker. Ook bedankt voor het feit dat jij (en Bonferonni) me soms corrigeerde als ik iets te enthousiast was.

Multicenter onderzoek is complex, maar waardevol. Daarom ook veel dank aan de medeonderzoekers van de PGS-studie in Groningen en Amsterdam. Maaike

van der Heide, Jacorina van Hoften, Paul Keating en Jorien Kerstjens, dank voor jullie hulp bij het onderzoek naar dysmorfieën bij de kinderen.

Moniek Twisk en Sebastiaan Mastenbroek, bedankt voor het gebruik van jullie databases en alvast gefeliciteerd met jullie promoties binnenkort!

Collega's van de afdeling Ontwikkelingsneurologie: Hylco Bouwstra, Cornill Blauw-Hospers, Corina de Jong, Lieke Peters, Tjitske Hielkema, Lieke van Balen, Tineke Dirks, Ineke Bakker, Victorine de Graaf-Peters, Leo van Eykeren en Michiel Schrier. Allen bedankt voor de betrokkenheid en de jarenlange goede samenwerking. Loes de Weerd, secretaresse van de ON, aan jou ben ik ook veel dank verschuldigd voor alle klussen die je voor het onderzoek op je hebt genomen en de administratieve hulp bij de laatste loodjes. In het bijzonder wil ik ook Saskia van Goor bedanken voor haar collegialiteit. Je zat een gang verderop en dat was het perfecte excuus om er af en toe even tussenuit te kunnen voor een 'bakkie'. We hebben inderdaad heel wat onderzoekers-lief-en-leed gedeeld; van peuterperikelen tot onze 'grote vrienden en vriendinnen' en enkele bijzondere associaties. Ik kon na een bezoekje aan jou altijd weer energiek en lachend verder.

Mijn perifere opleider en opleidingskliniek; Dr. D.J. Bekedam. Beste Dick, toevallig dat jouw wetenschappelijke carrière ook is begonnen bij de ON in het UMCG; leuk dat je komt opponeren. Ik wil jou, de vakgroep Gynaecologie en alle collega's van de O&G van het OLVG bedanken voor het geweldige leerklimaat, het vertrouwen en de bijzonder gastvrije plek die het 'Onze Lieve' echt is.

Mijn paranimfen, Maaike Haadsma en Kirsten Heineman. Eerst stond ik naast jullie, maar wat ben ik blij dat jullie nu ook naast mij kunnen en willen staan!

Lieve Maaike, het begon allemaal met lange lijsten te includeren patiënten, maar al snel volgden een heel goede samenwerking, uitermate gezellige congresbezoeken en vervolgens steeds meer terrasjes, avondjes en etentjes inclusief partners (en bitterballen!).

Lieve Kirsten, jij was eerst een collega bij de ontwikkelingsneurologie, maar werd ook al snel een vriendin. Je leerde me de praktische kneepjes van het onderzoeken van zuigelingen en wat was het fijn met jou een kamer te delen. Je geduld, humor en attentheid maakten het onderzoeksleven zoveel leuker!

Naast onderzoek is er nog zoveel meer! Ik wil deze gelegenheid ook graag gebruiken om alle mensen die dicht bij me staan te bedanken. Vrienden en vriendinnen van onder andere school, studie, volleybal, coschappen, werk. Allen bedankt voor de *'joie de vivre'*.

Lieve Pieter, Patricia, Marieke, Annemiek, Alinde en ook Fred, Tamara, Anouk, Joppe, Maike en Bastiaan. Broer, schoonfamilie, nichtjes, maar ook alle andere familieleden. Wat ben ik gelukkig met zo'n fijne, warme familie om me heen.

Lieve Lourens, grote broer. Jij excelleerde in de kunst van het genieten van het leven. Helaas veel korter dan de 100 jaren die je gepland had. We missen jou en al je (soms wat sterke) verhalen enorm. Je blijft een inspiratiebron voor me, omdat je altijd 'gewoon' hebt gedaan wat je wilde, ook al werd dit door anderen soms buitensporig of excentriek gevonden. Ik weet dat je trots op me was.

Lieve Pap en Mam, bedankt voor jullie grenzeloze liefde, steun, en vertrouwen. Jullie oneindige toewijding aan mijn opleiding is hartverwarmend, daarom draag ik graag dit proefschrift aan jullie op.

Merwin, mijn Lief. Met jou ben ik zo gelukkig als alle clichés beweren. Jij bent altijd zo positief, relativerend, grappig en ondernemend. Samen voel ik me enorm sterk. Dank voor wie je voor me bent. Ik hoop nog de prachtigste avonturen met jou te mogen beleven...

OVER DE AUTEUR

Karin Middelburg werd op 24 juli 1979 geboren in Assen. Zij volgde haar middelbare schoolopleiding aan het Dr. Nassau College te Assen, waar zij in 1997 haar VWO diploma behaalde. Een jaar later kon zij starten met de studie Geneeskunde aan de Rijksuniversiteit Groningen, waar in 2004 het arts-examen werd behaald. Vervolgens startte zij in 2005 onder leiding van prof. dr. M. Hadders-Algra, prof. dr. M.J. Heineman en prof. dr. A.F. Bos met het promotieonderzoek dat tot dit proefschrift heeft geleid. Tijdens het onderzoek in het Universitair Medisch Centrum Groningen begon zij met de opleiding ter registratie tot Epidemioloog B. In 2009 was ze werkzaam als arts-assistent Obstetrie en Gynaecologie in het Martini Ziekenhuis, Groningen. In april 2010 begon ze met de opleiding tot gynaecoloog in het Onze Lieve Vrouwe Gasthuis te Amsterdam (opleider dr. D.J. Bekedam), vanaf oktober 2011 zal ze de opleiding in het Academisch Medisch Centrum, Amsterdam vervolgen (opleider prof.dr. M.J. Heineman).

LIST OF PUBLICATIONS

Middelburg KJ, Heineman MJ, Bos AF, Hadders-algra M. Neuromotor, cognitive, language an behavioural outcome in children born following IVF or ICSI – a systematic review. *Hum Reprod Update* 2008; 14: 219-231.

Middelburg KJ, Heineman MJ, Bos AF, Pereboom M, Fidler V, Hadders-algra M. the Groningen ART cohort study: ovarian hyperstimulation and the in vitro procedure do not affect neurological outcome in infancy. *Hum Reprod* 2009; 24: 3119-26.

Middelburg KJ, Haadsma ML, Heineman MJ, Bos AF, Hadders-Algra M. Ovarian hyperstimulation and the in-vitro procedure do not influence early neuromotor development, a history of subfertility does. *Fertil Steril* 2010; 93: 544-553.

Hadders-Algra M, Heineman KR, Bos AF, **Middelburg KJ**. The assessment of minor neurological dysfunction in infancy using the Touwen Infant Neurological Examination: strengths and limitations. *Dev Med Child Neurol*. 2010; 52: 87-92.

Pelinck MJ, Keizer M, Hoek A, Simons AHM, Schelling K, **Middelburg KJ**, Heineman MJ. Perinatal outcome in singletons after modified natural cycle IVF and standard IVF with ovarian stimulation. *Eur J Obstet Gynecol Reprod Biol*, 2010; 148: 56-61.

Heineman KR, **Middelburg KJ**, Hadders-algra M. Development of adaptive motor behaviour in typically developing infants. *Acta Paediatr*. 2010; 99: 618-624.

Middelburg KJ, Heineman MJ, Haadsma ML, Bos AF, Kok JH, Hadders-Algra M. Neurological condition of children born following preimplantation genetic screening. *Pediatr Res*. 2010; 67: 430-434.

Kikkert HK, **Middelburg KJ**, Hadders-Algra M. Maternal anxiety is related to infant neurological condition, paternal anxiety is not. *Early Hum Dev*. 2010; 86: 171-177.

Heineman KR, la Bastide-van Gemert S, Fidler V, **Middelburg KJ**, Bos AF, Hadders-Algra M. Construct validity of the Infant Motor Profile: Relation with prenatal, perinatal and neonatal risk factors. *Dev Med Child Neurol*. 2010; 52: e209-215.

Pelinck MJ, Hadders-Algra M, Haadsma M, Nijhuis W, Kiewiet S, Hoek A, Heineman MJ, **Middelburg KJ**. Is birthweight of singletons born after IVF reduced by ovarian stimulation? *Reprod Biomed Online*. 2010; 21: 245-251.

Schendelaar P, **Middelburg KJ**, Bos AF, Heineman MJ, Jongbloed-Pereboom M, Hadders-Algra M. The Groningen ART cohort study: the effects of ovarian hyperstimulation and the IVF laboratory procedures on neurological condition at 2 years. *Hum Reprod*. 2011; 26: 703-712.

Jongbloed-Pereboom M, **Middelburg KJ**, Heineman MJ, Bos AF, Hadders-Algra M. The Groningen assisted reproductive technologies cohort study: developmental status and behavior at 2 years. *Fertil Steril*. 2011; 95: 2283-2289.

Middelburg KJ, van der Heide M, Houtzager B, Jongbloed-Pereboom M, Fidler V, Bos AF, Kok J, Hadders-Algra M. Mental, psychomotor, neurological, and behavioral outcome of 2-year-old children born following preimplantation genetic screening: follow-up of a randomized controlled trial. *Fertil Steril*, 2011; 96: 165-169.

van der Heide M, Beukers F, **Middelburg KJ**, Cobben JM, Mastenbroek S, Breur R, van der Lee JH, Hadders-Algra M, Bos AF, Kok JH. Morphological abnormalities in 2-year-old children born after IVF/ICSI with Preimplantation Genetic Screening. Follow-up of a randomised controlled trial. submitted.

Beukers F, Houtzager BA, Paap MCS, **Middelburg KJ**, Hadders-Algra M, Bos AF, Kok JH. Parental psychological distress and anxiety after a successful IVF/ICSI procedure with and without Preimplantation Genetic Screening. Follow-up of a randomised controlled trial. submitted.

Seggers J, Haadsma ML, Bos AF, Heineman MJ, Keating P, **Middelburg KJ**, van Hoften JC, Veenstra-Knol HE, Kok JH, Cobben JM, Hadders-Algra, M. The influence of ovarian hyperstimulation and the IVF laboratory procedure on the prevalence of dysmorphic features. submitted.

Jongbloed-Pereboom M, **Middelburg KJ**, Heineman MJ, Bos AF, Haadsma, ML, Hadders-Algra M. The impact of assisted reproductive technology on parental well-being and anxiety one year after childbirth. submitted.