

University of Groningen

## Neurodevelopmental outcome of children born following assisted reproductive technology

Middelburg, Karin Janette

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*

Publisher's PDF, also known as Version of record

*Publication date:*

2011

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Middelburg, K. J. (2011). *Neurodevelopmental outcome of children born following assisted reproductive technology: 0-2 years*. s.n.

### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

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

**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 ( $\kappa$  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 (Prechtel, 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:  $-\ln(59.5 - \text{NOS})$ , and the fluency score was transformed into:  $-\ln(14.5 - \text{fluency score})$ . 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  $(14.5 - \text{medB}) / (14.5 - \text{medA})$ , 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 I - DEMOGRAPHIC CHARACTERISTICS

Characteristics	COH-IVF <sup>a</sup> n = 68	MNC-IVF <sup>a</sup> n = 57	NC <sup>a</sup> n = 90
Gender: male, n (%)	36 (53%)	27 (47%)	46 (51%)
First born, n (%)	47 (69%)	38 (67%)	55 (61%)
<b>Birth characteristics:</b>			
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 <sup>b</sup> ; 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 <sup>c</sup> , n (%)	20 (29%)	16 (28%)*	40 (44%)*
<b>Neonatal characteristicse:</b>			
Apgar score 5 min < 7, n (%)	0 (0%)	0 (0%)	1 (1%)
Neonatal intensive-care admission, n (%)	1 (2%)	2 (4%)	7 (8%)
<b>Parental characteristicse:</b>			
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 <sup>d</sup> , n (%)	22 (32%)	22 (39%)	41 (46%)
Education level father high <sup>d</sup> , n (%)	29 (45%)	19 (34%)	33 (37%)
<b>Fertility parameterse:</b>			
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%)**
<b>Corrected age at examination:</b>			
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  $\chi^2$  or Fisher's exact test were used to compare groups; \* P < .05, \*\* P < .017 (Bonferonni correction), \*\*\* P < .001.

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

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

<sup>c</sup> Signs of fetal distress denoted by meconium stained amniotic fluid and/or cardiotocographic signs and/or acidosis.

<sup>d</sup> University education or vocational colleges.

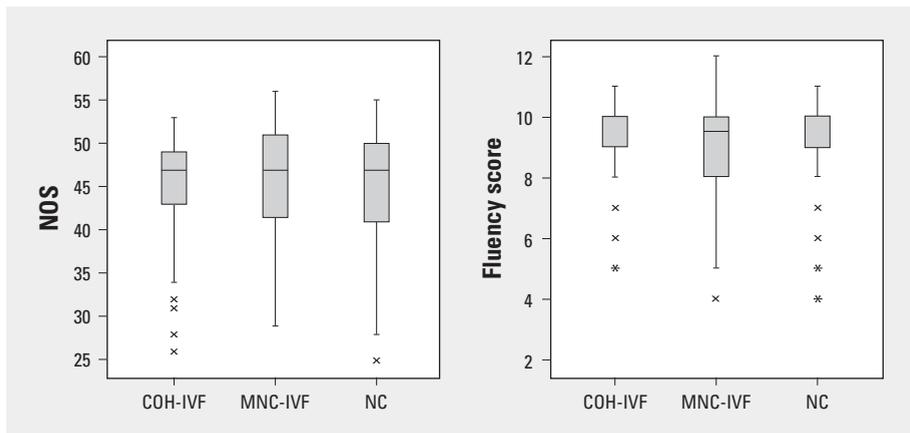
<sup>e</sup> Missing values for variables in the COH-IVF, MNC-IVF and NC group: < 3, < 2, and < 4, respectively.

### **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.**

<b>Outcome measure: Neurological optimality score (-ln (59.5 - NOS))</b>		
<b>compared groups</b>	<b>adjusted median difference [95% CI]<sup>e</sup></b>	<b>p-value</b>
COH-IVF vs. MNC-IVF	-0.066 [-0.226, 0.094] <sup>b</sup>	0.415
MNC-IVF vs. NC	0.067 [-0.084, 0.218] <sup>c</sup>	0.385
COH-IVF vs. NC	-0.004 [-0.155, 0.146] <sup>d</sup>	0.957
<b>Outcome measure: Fluency score (-ln (14.5 - Fluency score))</b>		
<b>compared groups</b>	<b>adjusted median difference [95% CI]<sup>e</sup></b>	<b>p-value</b>
COH-IVF vs. MNC-IVF	0.049 [-0.055, 0.152] <sup>b</sup>	0.354
MNC-IVF vs. NC	-0.010 [-0.116, 0.095] <sup>c</sup>	0.846
COH-IVF vs. NC	0.038 [-0.065, 0.142] <sup>d</sup>	0.464
<b>Outcome measure: Complex MND<sup>a</sup></b>		
<b>compared groups</b>	<b>adjusted odds ratio [95% CI]<sup>f</sup></b>	<b>p-value</b>
COH-IVF vs. MNC-IVF	1.30 [0.38, 4.42] <sup>b</sup>	0.670
MNC-IVF vs. NC	1.44 [0.40, 5.25] <sup>c</sup>	0.576
COH-IVF vs. NC	1.85 [0.55, 6.24] <sup>d</sup>	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.

<sup>a</sup> MND: Minor Neurological Dysfunction  
<sup>b</sup> Adjusted for gestational age and vanishing twins.  
<sup>c</sup> Adjusted for gestational age, signs of fetal distress and time to pregnancy.  
<sup>d</sup> Adjusted for gestational age, birth weight, vanishing twins and time to pregnancy.  
<sup>e</sup> Multivariate linear regression was used for NOS and fluency score.  
<sup>f</sup> 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 <sup>a</sup> <i>n</i> = 68	MNC-IVF <sup>a</sup> <i>n</i> = 57	NC <sup>a</sup> <i>n</i> = 90
<b>Neurological outcome at 4 months</b>			
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
<b>Neurological outcome at 10 months</b>			
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
<b>Neurological outcome at 18 months</b>			
Normal	44 (67%)	40 (71%)	69 (79%)
Simple MND <sup>b</sup>	15 (23%)	11 (20%)	12 (14%)
Complex MND <sup>b</sup>	7 (11%)	5 (9%)	6 (7%)
Total number of children tested	66	56	87
<b>Clusters of dysfunction at 18 months</b>			
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,  $\chi^2$  or Fisher's Exact test.

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

<sup>b</sup> 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 <sup>b</sup>	COH-IVF <sup>a</sup> <i>n</i> = 66	MNC-IVF <sup>a</sup> <i>n</i> = 56	NC <sup>a</sup> <i>n</i> = 84
Neurological classification at 4, 10, and 18 months			
normal, normal, normal	56 (85%)	49 (88%)	77 (92%)
MND, normal, normal <sup>c</sup>	2 (3%)	2 (4%)	1 (1%)
normal, MND, normal <sup>c</sup>	1 (2%)	0 (0%)	0 (0%)
MND, MND, normal <sup>c</sup>	0 (0%)	0 (0%)	0 (0%)
normal, normal, C-MND <sup>d</sup>	3 (4%)	5 (9%)	3 (4%)
MND, normal, C-MND <sup>d</sup>	2 (3%)	0 (0%)	2 (2%)
normal, MND, C-MND <sup>d</sup>	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.

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

<sup>b</sup> Only children who were assessed at all ages were used in this analysis. MND: minor neurological dysfunction, C-MND: complex - minor neurological dysfunction.

<sup>c</sup> Improvement in developmental trajectory with age.

<sup>d</sup> 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 from 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.

