# Neurodevelopmental outcome of children born following assisted reproductive technology

Middelburg, Karin Janette

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 $\chi^2$ 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 I - 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 |

▶▶▶

6

**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