Neurodevelopmental outcome of children born following assisted reproductive technology

Middelburg, Karin Janette

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

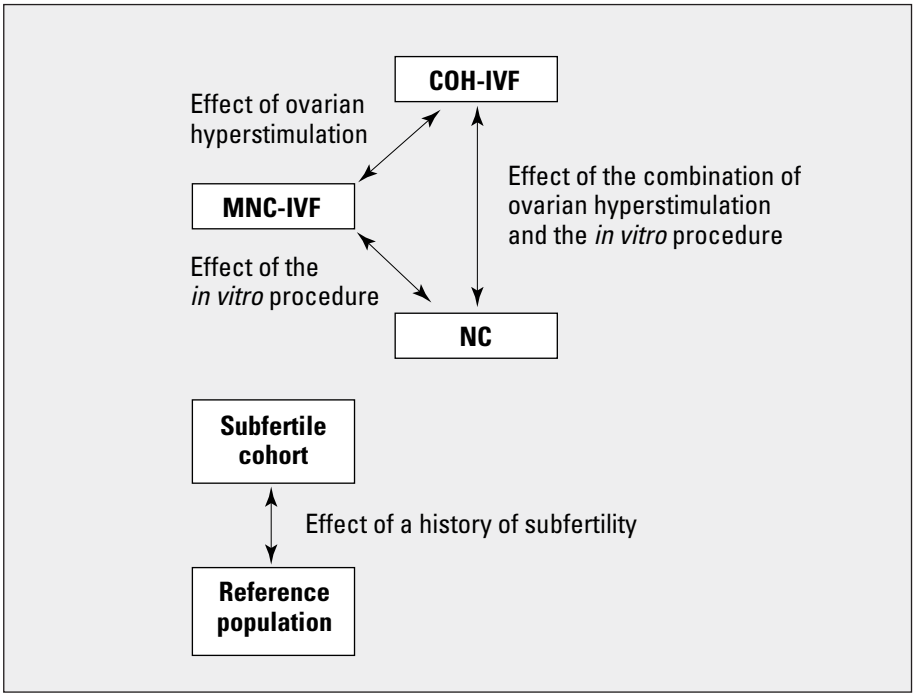
Nowadays a considerable number of children are born after IVF or intracytoplasmic sperm injection (ICSI) (Wright *et al.*, 2008;Andersen *et al.*, 2008), and as a result of these large numbers even subtle changes in the health of these children are of importance to society. It is well known that adverse 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.