

Cognitive and behavioural outcome of children born after IVF at age 9 years

K.R. Heineman^{1,2,*}, D.B. Kuiper¹, S.la Bastide-van Gemert³,
M.J. Heineman⁴, and M. Hadders-Algra¹

¹Department of Paediatrics, Division of Developmental Neurology, University of Groningen, University Medical Center Groningen, Hanzeplein 1, Groningen 9713 GZ, The Netherlands ²SEIN, Stichting Epilepsie Instellingen Nederland, Dr Denekampweg 20, 8025 BV Zwolle, The Netherlands ³Department of Epidemiology, University of Groningen, University Medical Center Groningen, Hanzeplein 1, Groningen 9713 GZ, The Netherlands ⁴Department of Obstetrics and Gynaecology, Academic Medical Center, Amsterdam University Medical Centre, Meibergdreef 9, Amsterdam 1105 AZ, The Netherlands.

*Correspondence address. Department of Paediatrics, Division of Developmental Neurology, University Medical Center Groningen, Hanzeplein 1 9713 GZ Groningen, The Netherlands. Tel: +31 620625570; Fax: +31 50 3619158; E-mail: k.r.heineman@umcg.nl

Submitted on March 7, 2019; resubmitted on July 15, 2019; editorial decision on August 29, 2019

STUDY QUESTION: Do ovarian stimulation (OS) and the *in vitro* laboratory procedures affect offsprings' cognitive and behavioural outcome at 9 years?

SUMMARY ANSWER: OS and the *in vitro* laboratory procedures or the combination of both were not associated with cognitive and behavioural outcome at age 9 years.

WHAT IS KNOWN ALREADY: ART is not associated with an adverse short-term developmental outcome of the offspring, but limited knowledge is available on the offspring's long-term neurodevelopmental condition.

STUDY DESIGN, SIZE, DURATION: A 9-year longitudinal, assessor-blinded, prospective follow-up study of 169 out of 215 singletons (79%) born between March 2005 and December 2006 was performed.

PARTICIPANTS/MATERIALS, SETTING, METHODS: Singletons born following IVF or ICSI with OS ($n = 57$), born after modified natural cycle IVF/ICSI (MNC-IVF/ICSI; $n = 46$) and born after natural conception to subfertile couples (Sub-NC; $n = 66$), were assessed at 9 years. This study design, with two ART groups and a subfertile reference group, allows for disentangling the effects of OS and ART procedures on developmental outcome. Cognitive outcome was evaluated with the Wechsler abbreviated scale of intelligence and the NEPSY-II. Behaviour was assessed with the child behaviour checklist (CBCL) and teacher report form (TRF). Univariable analyses and multiple linear regression models were used.

MAIN RESULTS AND THE ROLE OF CHANCE: There was no significant difference in intelligence quotient (IQ) scores between ART groups (mean IQ (95% CI): OS 114.8 (83.2–142.6); MNC 114.0 (90.2–140.8); Sub-NC 115.4 (87.9–141.2), $P = 0.746$). Multivariable analyses did not reveal a statistically significant association between ART group and total, verbal and performance IQ. CBCL and TRF scores did not differ significantly between ART groups ($P = 0.090$ and 0.507 , respectively). Multivariable analyses did not demonstrate a statistically significant association between ART group and CBCL and TRF total, or internalising and externalising T-scores. No significant correlations between time to pregnancy (TTP)—a proxy for the severity of parental subfertility—and outcome measures were found (Spearman rho between -0.050 and 0.049 , NS), which was confirmed with multivariable analyses.

LIMITATIONS, REASONS FOR CAUTION: The attrition rate of 21% may be considered as a limitation of the study; however, after a follow-up period of 9 years, this rate is generally considered acceptable, and there were no significant differences in background characteristics between children with and without follow-up, making an attrition-related selection bias less likely. Another limitation of the study is the relatively small sample size, which could contribute to selection bias, hamper generalizability to the ART population and lead to false negative findings as a result of underpowering. An *a priori* power analysis on total IQ indicated that the OS-IVF/ICSI and Sub-NC groups should contain 64 children, confirming that our study including 57 and 66 children, respectively, was slightly underpowered.

WIDER IMPLICATIONS OF THE FINDINGS: Our study indicated that OS and the *in vitro* laboratory procedures or the combination of both and TTP were not associated with cognitive and behavioural outcome at 9 years. These are reassuring results for both parents and clinicians involved in ART.

STUDY FUNDING/COMPETING INTEREST(S): The study was financially supported by the University Medical Center Groningen (UMCG), two graduate schools of the UMCG (BCN and SHARE) and the Cornelia Stichting. The sponsors of the study had no role

in study design, data collection, data analysis, data interpretation or writing of the report. The authors have no conflicts of interest to declare.

Key words: IVF / school-age / cognitive outcome / behavioural outcome / ovarian stimulation / developmental outcome / intelligence quotient

Introduction

In The Netherlands, currently almost 3% of the children are born with the help of ART, such as IVF or ICSI, and this percentage is increasing (websites CBS and NVOG). This raises the important question of whether ART influences developmental outcome of these children. Factors that could play a role are hormonal influences such as ovarian stimulation (OS), effects of the laboratory procedures used in IVF and ICSI and possible influences of vanishing twins (Oliviennes, 1993; Pinborg, 2005; Kapiteijn, 2006). Also, the subfertility of the parents could be an influencing factor in developmental outcome (Schendelaar et al., 2014, 2016).

Despite the higher rate of unfavourable perinatal outcomes, especially in terms of preterm birth and low birthweight (Helmerhorst et al., 2004), ART is not associated with adverse effects on developmental outcome up to and including 5 years of age (Middelburg et al., 2008; Wagenaar et al., 2008; Carson et al., 2011; Hart and Norman, 2013b; Fauser et al., 2014). Some metabolic differences such as higher blood pressure, higher fat under the skin and higher fasting glucose concentrations have been found (Hart and Norman, 2013a; Fauser et al., 2014). However, at present, it is not clear whether there are long-term developmental and health differences between children born after ART and their peers born from fertile parents. With respect to cognitive development beyond the age of 5 years, a limited number of studies are available with conflicting evidence: intelligence quotient (IQ) at school age in children born after ART was similar or somewhat higher than that of children born to fertile parents (Leunens et al., 2006; Leunens et al., 2008; Barbuscia and Mills, 2017), executive functions in adolescence were similar in children born with or without ART (Wagenaar et al., 2009b), and however, academic performance scores in adolescence were slightly lower in ART children (Spangmose et al., 2017). For behavioural development, the available literature also does not report consistent results (Wagenaar et al., 2008): some studies found that ART children showed more signs of behavioural problems compared with children born to fertile parents, but the type of behavioural problems was inconsistent and considered not to be in a clinically relevant range (Wagenaar et al., 2009a; Wagenaar et al., 2011; Carson et al., 2013; Barbuscia et al., 2019). As especially cognitive functions show major developmental changes after the age of 5 years, it is conceivable that ART components may affect these functions at school age. Therefore, we assessed cognition and behaviour of the offspring of the Groningen assisted reproductive technologies (ART) cohort study at 9 years of age. The ART study is a longitudinal follow-up study on the development of children born after ART (Middelburg et al., 2009), which was designed to disentangle the effects of OS and the *in vitro* procedure itself on developmental outcome. The study cohort consists of children born following IVF or ICSI with conventional ovarian stimulation (OS-IVF/ICSI), children born following IVF/ICSI in the modified natural cycle (MNC-IVF/ICSI) and children who were born after natural conception to subfertile couples (Sub-NC group).

Previous studies on the Groningen ART cohort did not find differences in neurological, cognitive and behavioural development up to the age of 4 years between these three ART groups (Middelburg et al., 2009; Schendelaar et al., 2014, 2016). However, severity of subfertility was associated with suboptimal neurological and cognitive development at the age of 4 years (Schendelaar et al., 2014, 2016) and with neurological outcome at age 9 years (Drenth Olivares et al., 2019).

The aim of the current ART study was to investigate the effects of OS and the *in vitro* procedure itself on cognitive and behavioural outcome at the age of 9 years. The study design enables us to disentangle possible effects of OS and the *in vitro* laboratory procedures, as both an ART group with OS (OS-IVF/ICSI) and an ART group without OS (MNC-IVF/ICSI) are included. In addition, the effect of time to pregnancy (TTP)—a proxy for the severity of subfertility—on developmental outcome was examined. Cognitive outcome was assessed by means of the Wechsler abbreviated scale of intelligence (WASI) (Wechsler, 1999) and the NEPSY (Korkman et al., 2007a), and behavioural outcome was assessed with the child behaviour checklist (CBCL) (Achenbach and Rescoria, 2001) and teacher report form (TRF) (Achenbach and Rescoria, 2001).

Materials and Methods

Ethical approval

The Medical Ethical Committee of the University Medical Center Groningen (UMCG) approved the study design. Parents gave written informed consent.

Participants

The participants of the Groningen ART study were recruited prenatally at the Department of Reproductive Medicine of the UMCG. Pregnant couples who achieved a singleton pregnancy via IVF or ICSI with a term date between March 2005 and December 2006 were eligible (Middelburg et al., 2010). This could be either OS-IVF/ICSI with conventional hormonal OS or MNC-IVF/ICSI in which use of hormones was minimal. Details on the procedures in MNC-IVF/ICSI have been described previously (Pelinck et al., 2007, 2008). Couples treated with cryopreserved or donated oocytes or embryos were not suitable for inclusion. A subfertile natural conception (Sub-NC) control group was formed by inviting all couples who conceived naturally while on the waiting list for fertility evaluation or treatment during the study period. These couples had been trying to conceive naturally for at least 1 year and were then referred to the Department of Reproductive Medicine. All couples were approached during the third trimester of pregnancy. Prenatal, perinatal and demographic information was collected from the parents and from medical records with standardised charts during the first follow-up assessment of the ART study at 2 weeks corrected age (Middelburg et al., 2010).

Neonatally, 68 infants were included in the OS-IVF/ICSI group, 57 in the MNC-IVF/ICSI group and 90 in the Sub-NC group (Middelburg *et al.*, 2010). At 9 years, 57 OS-IVF/ICSI (84%), 46 MNC-IVF/ICSI (81%) and 66 Sub-NC (73%) children were assessed (Supplementary Fig. S1). Reasons for attrition were withdrawal of consent due to overloaded family agendas, inability to contact families and emigration. There were no significant differences in all background characteristics (Table I) between the children who had and who did not have follow-up at the age of 9 years (Mann–Whitney U test or Fisher exact test, *P*-values ranging from 0.155 to 1.000).

Table I shows the background characteristics of the participating children and their parents. Most background characteristics did not differ between the ART groups, except TTP which was shortest in the Sub-NC group (OS-IVF/ICSI versus Sub-NC *P* < 0.001; MNC-IVF/ICSI versus Sub-NC *P* = 0.003), the number of pregnancies with vanishing twins which was higher in the OS-IVF/ICSI group compared with the Sub-NC group (*P* = 0.009), gestational age and birthweight which were lower in the OS IVF/ICSI group than in the Sub-NC group (*P*-values, respectively, 0.036 and 0.049) and paternal age at conception which was higher in the OS-IVF/ICSI group compared with the MNC IVF/ICSI group (*P* = 0.020).

Cognitive assessment

To assess cognitive outcome in terms of IQ at age 9 years, the WASI (Wechsler, 1999) was used. This condensed IQ test yields three IQ scores: verbal IQ score with subtests vocabulary and similarities; performance IQ (PIQ) with subtests block design and matrices; and full-scale IQ based on all four subtests. The WASI is a validated screening IQ test and is suitable for ages 6–89 years (Wechsler, 1999). The WASI has good correlations with more extensive IQ tests such as the Wechsler Adult Intelligence Scale-III and the Wechsler Intelligence Scale for Children-III (Wechsler, 1999).

In order to assess specific domains of neuropsychological development, we used the Dutch version of the NEPSY-II (Korkman *et al.*, 2007a, b). The NEPSY-II is designed for children aged 5–12 years and has good psychometric properties (Korkman *et al.*, 2007b). In the current study, three of the six NEPSY-domains were assessed: attention and executive functions; memory and learning; and social cognition, as previous studies suggested that ART-offspring may have an increased risk of impairments in these domains (Källén *et al.*, 2011; Knoester *et al.*, 2007). The domain attention and executive functions consist of two subtests, namely inhibition and design fluency, and yield 14 scores. The domain memory and learning comprise three subtests (memory for names, memory for names delayed and narrative memory) and generate six scores. The domain social cognition comprises two subtests: affect recognition and theory of mind, and yields three scores. All NEPSY scores of the various subtests were dichotomized into typical or atypical (classifications weak or very weak) according to instructions in the manual: scaled scores between 4 and 5, and percentile scores between 3 and 10 are considered as weak, scaled scores below 4 and percentile scores below 3 are considered as very weak (Korkman *et al.*, 2007a, b). The number of atypical scores for each NEPSY domain was calculated.

Behavioural assessment

Behavioural outcome was assessed by means of the CBCL and the TRF (Achenbach and Rescoria, 2001). Reliability and internal consistency of both CBCL and TRF are good (Achenbach and Rescoria, 2001). The CBCL and TRF are questionnaires on the behaviour of the child, which are filled in by one of the parents or the teacher, respectively. They both include 113 items on behavioural and emotional problems, which are all scored on a three-point scale: not true, somewhat or sometimes true and very true or very often true. The sum of all items yields a total problem score, an internalising problem score with domains emotionally reactive, anxious or depressed, somatic complaints and withdrawn, and an externalising problem score with domains attention problems and aggressive behaviour. The raw scores are transformed into T-scores, where higher T-scores stand for more problematic behaviour. T-scores above 63 are in the clinical range, between 60 and 63 are considered borderline and below 60 are normal. We dichotomized scores into typical (normal) and atypical (borderline and clinical). If information on more than eight of the 113 items was missing, we excluded the questionnaire.

Statistical analyses

Differences between the groups in background characteristics and outcome parameters were first tested with univariable statistics, using parametric and non-parametric tests when appropriate. Next, multiple linear regression with backward elimination was performed. Possible confounders (gender, ICSI, smoking during pregnancy, alcohol use during pregnancy, use of folic acid, vanishing twins, gestational age, preterm birth, birthweight, low birthweight, small for gestational age, Caesarean section, signs of foetal distress, maternal age, paternal age, educational level of the mother, educational level of the father and causes of subfertility; see Table I) were entered into these analyses if they had an effect on the outcome variable at the significance level of 0.25 in the univariate analyses. For each of the outcome parameters (WASI, NEPSY, CBCL and TRF), two models were created: one with and one without the categorical variable ART group in addition to the set of confounders, in order to see whether ART group was a significant factor. TTP was defined as the time in decimal years between the start of unprotected intercourse or the end of a previous pregnancy and conception. Spearman rho correlations between TTP and outcome measures were calculated. In addition, multivariable analyses of associations between TTP and outcome parameters were performed with the same set of possible confounders as above, excluding ART group. Throughout the analyses, probability values of 5% or less were considered statistically significant. SPSS Statistics version 25.0 (SPSS Inc., Chicago, IL, USA) was used.

Results

Effects of ART

There were no differences in total IQ between the three ART groups with a mean total IQ for the OS-IVF/ICSI group of 114.8 (range 79–150), for the MNC-IVF/ICSI group of 113.8 (range 77–150) and for the Sub-NC group of 115.1 (range 86–147). Also, no significant differences were found between the ART groups for verbal and performance IQ (Table II). Furthermore, neither the number of

Table 1 Characteristics of participating children and parents of the Groningen ART cohort study groups.

Characteristics	OS-IVF/ICSI n = 57	MNC-IVF/ICSI n = 46	Sub-NC n = 66
Child characteristics			
Male gender, n (%)	32 (56)	22 (48)	33 (50)
First born, n (%)	38 (67)	33 (72)	39 (59)
Age at examination in months, median (range)	110.4 (108.5–126.5)	110.3 (108.2–131.8)	109.9 (100.7–119.9)
Fertility parameters			
TTP in years ^a , median (range)	4.0 (0.1–13.3)***	3.8 (0.1–7.5)**	2.0 (0.1–11.3)****/**
ICSI, n (%)	37 (65)	21 (46)	n.a.
Gestational characteristics			
Smoking during pregnancy, n (%)	6 (11)	5 (11)	6 (9)
Alcohol use during pregnancy, n (%)	3 (5)	0 (0)	2 (3)
Use of folic acid during pregnancy ^b , n (%)	50 (93)	46 (100)*	57 (86)*
Vanishing twins, n (%)	6 (11)**	1 (2)	0 (0)**
Birth characteristics			
Gestational age in weeks, median (range)	39.4 (33.4–42.3)*	39.9 (34.6–42.6)	40.1 (30.1–42.6)*
Preterm birth (<37 weeks), n (%)	6 (11)	6 (13)	4 (6)
Birthweight in grams, mean (σ)	3340 (563)*	3382 (604)	3594 (517)*
Low birthweight, n (%)	3 (5)	4 (9)	2 (3)
Small-for-gestational age, n (%)	0 (0)	3 (7)	1 (2)
Caesarean section, n (%)	15 (26)	8 (17)	19 (29)
Neonatal characteristics			
NICU admission, n (%)	1 (2)	2 (4)	4 (6)
Apgar score at 5 min <7 ^b , n (%)	0 (0)	0 (0)	0 (0)
Breastfed for >6 weeks ^b , n (%)	29 (52)	21 (46)	33 (51)
Signs of foetal distress ^d , n (%)	19 (33)	12 (26)	27 (41)
Parental characteristics			
Maternal age at conception, median (range)	33.2 (27.0–40.9)	32.8 (26.2–37.5)	33.7 (23.1–40.3)
Paternal age at conception ^b , median (range)	36.4 (27.5–56.1)*	33.7 (28.3–47.8)*	35.4 (25.5–48.7)
Educational level mother high ^e , n (%)	20 (35)	20 (43)	31 (47)
Educational level father high ^{b,e} , n (%)	26 (48)	15 (33)	25 (38)
Cause of subfertility^f			
Primary subfertility, n (%)	31 (54)	29 (63)	32 (49)
Tubal pathology, n (%)	11 (19)	9 (20)	5 (8)
Male factor, n (%)	32 (56)**	25 (54)*	19 (29)**/*
Other causes, n (%)	18 (32)*	6 (13)*	12 (18)
Unknown cause, n (%)	7 (12)***	8 (17)***	35 (53)****/****

OS-IVF/ICSI: children born following ovarian stimulation IVF or ICSI, MNC-IVF/ICSI: children born following modified natural cycle IVF or ICSI, Sub-NC: naturally conceived children born to subfertile parents, NICU: neonatal intensive care unit.

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$ Student's *t*-tests, Fisher's exact tests and Mann–Whitney U-tests were performed to investigate differences between the groups.

^aTime to pregnancy (TTP) of the three ART groups was recorded in years and months and finally converted into decimal years. In case of a miscarriage, the onset of TTP was reset; therefore, TTP may be shorter than 1 year.

^bMissing data in three groups: Use of folic acid—OS-IVF: $n = 3$; Apgar score at 5 min <7—MNC-IVF: $n = 1$, OS-IVF: $n = 1$; breastfed >6 weeks—OS-IVF: $n = 1$, Sub-NC: $n = 1$; paternal age at conception—OS-IVF: $n = 2$, MNC-IVF: $n = 1$; educational level father—OS-IVF: $n = 3$, MNC-IVF: $n = 1$.

^cBirthweight for gestational age is below 2 SDs compared to a Dutch reference population.

^dSigns of foetal distress defined by meconium stained amniotic fluid and/or cardiotocographic signs and/or acidosis.

^eUniversity education or vocational colleges.

^fCouples may have more than one cause of subfertility; therefore, totals may exceed 100%.

Table II Cognitive and behavioural outcomes at age 9 years of the Groningen ART cohort study groups.

	OS-IVF/ICSI (n = 57)	MNC-IVF/ICSI (n = 46)	Sub-NC (n = 66)	P-value*
WASI				
Total IQ, mean (95% CI)	114.7 (83.2–142.6)	113.8 (90.2–140.8)	115.1 (87.9–141.2)	0.746
Verbal IQ, mean (95% CI)	114.4 (84.2–139.8)	113.0 (89.0–144.5)	116.7 (87.9–145.3)	0.428
Performance IQ, mean (95% CI)	112.2 (86.8–139.0)	112.4 (89.0–137.8)	110.8 (88.7–139.6)	0.806
Total IQ < 80, n (%)	1 (2%)	1 (3%)	0 (0%)	0.515
Verbal IQ < 80, n (%)	1 (2%)	1 (3%)	0 (0%)	0.515
Performance IQ < 80, n (%)	1 (2%)	0 (0%)	1 (2%)	0.667
NEPSY				
NEPSY total number of atypical scores, median (range)	2.0 (0–10)	2.50 (0–11)	3.00 (0–9)	0.347
NEPSY total number of atypical scores on domain attention and executive function, median (range)	1.50 (0–10)	2.00 (0–7)	2.00 (0–6)	0.651
NEPSY total number of atypical scores on domain memory and learning, median (range)	0.00 (0–4)	0.00 (0–3)	0.00 (0–5)	0.677
NEPSY total number of atypical scores on domain social cognition, median (range)	0.00 (0–4)	0.00 (0–3)	0.00 (0–5)	0.677
CBCL				
CBCL T score on total behaviour problems, mean (SD)	51.0 (10.58)	45.9 (10.81)	49.1 (9.54)	0.090
CBCL, T score internalising problems, mean (SD)	52.1 (10.53)	46.5 (9.39)	50.7 (10.56)	0.119
CBCL, T score externalising problems, mean (SD)	48.4 (9.95)	44.7 (10.76)	47.6 (9.62)	0.167
CBCL atypical T score total behavioural problems, n (%)	14 (25%)	4 (9%)	10 (16%)	0.116
CBCL atypical T score internalising problems, n (%)	14 (25%)	3 (7%)	13 (21%)	0.061
CBCL atypical T score externalising problems, n (%)	9 (16%)	3 (7%)	8 (13%)	0.389
TRF				
TRF T score total behaviour problems, mean (SD)	49.6 (7.89)	48.6 (8.22)	50.4 (7.79)	0.507
TRF T score internalising problems, mean (SD)	51.1 (10.48)	49.6 (8.48)	49.4 (9.01)	0.760
TRF T score externalising problems, mean (SD)	47.0 (6.32)	48.9 (7.07)	50.1 (7.41)	0.072
TRF atypical T score total behavioural problems, n (%)	4 (9%)	4 (10%)	8 (14%)	0.722
TRF atypical T score internalising problems, n (%)	9 (20%)	6 (14%)	7 (12%)	0.477
TRF atypical T score externalising problems, n (%)	2 (5%)	3 (7%)	6 (10%)	0.562

CBCL: Child Behaviour Checklist, TRF: Teacher Report Form, WASI: Wechsler Abbreviated Scale of Intelligence, IQ: intelligence quotient, n: number, *tested with Kruskal–Wallis test

children with an IQ (total, verbal or performance) below 80 was significantly different between the three groups, nor the number of atypical NEPSY scores (Table II). Multivariable analyses confirmed that ART group was not associated with any of the cognitive outcome parameters.

Univariate analyses did not show significant differences in CBCL T-scores on total behaviour problems between ART groups (Table II). This was also the case for CBCL T-scores on internalising and externalising problems. In addition, no significant differences were found in the number of atypical CBCL T-scores between groups (Table II). This pattern was similar for the TRF T-scores (Table II). Multivariable analyses confirmed that there were no significant relations between ART and behavioural outcome parameters in terms of CBCL and TRF, except for the CBCL internalising T-score, where the MNC-IVF/ICSI group scored significantly lower (i.e. behaviourally better) than the OS-IVF/ICSI and the Sub-NC group (adjusted regression coefficient (95% CI): OS-IVF/ICSI versus MNC-IVF/ICSI 6.36 (2.19–10.53), MNC-IVF/ICSI versus Sub-NC –4.76 (–8.74 to –0.77)).

Effects of TTP

There was no significant correlation between TTP and total, verbal and performance IQ (Spearman rho –0.096, –0.057 and –0.096, respectively; NS). TTP was also not significantly correlated to the number of atypical NEPSY scores (Spearman rho 0.041, $P = 0.630$). Univariable statistics also indicated that TTP was not associated with the behavioural outcomes in terms of CBCL and TRF scores (Spearman rho correlations ranging from –0.050 to 0.049, NS). Multivariable analyses confirmed that TTP was not associated with any of the cognitive and behavioural outcome parameters (Table III).

Discussion

The current study did not find an effect of ART, i.e. the OS, the *in vitro* laboratory procedures or the combination of both, on cognitive and behavioural outcome in children at age 9 years. Assessments were performed with WASI (total, verbal and performance IQ), NEPSY-II, CBCL and TRF. The only exception to this was the finding of slightly better CBCL internalising scores in favour of the MNC-IVF/ICSI group

Table III Multiple regression analyses of the effect of OS, the *in vitro* procedure, a combination of both and TTP on IQ, NEPSY, CBCL and TRF.

Outcome variable	Adjusted regression coefficient (95% CI)			
	OS-IVF/ICSI versus MNC IVF/ICSI	MNC-IVF/ICSI versus Sub-NC	OS-IVF/ICSI versus Sub-NC	TTP (in the pooled subfertile group)
Total IQ ^a	2.57 (−3.54–8.67)	−1.37 (−7.17–4.42)	1.19 (−4.42–6.81)	−0.88 (−1.71–0.50)
Verbal IQ ^b	3.93 (−2.71–10.58)	−3.71 (−10.02–2.60)	0.22 (−5.90–6.34)	−0.43 (−1.54–0.88)
Performance IQ	0.78 (−5.69–7.25)	1.21 (−5.00–7.41)	1.39 (−4.51–7.29)	−1.42 (−2.01–0.33)
NEPSY ^c	−0.78 (−1.82–0.26)	−0.17 (−1.14–0.80)	−0.95 (−1.91–0.01)	−0.04 (−0.23–0.15)
CBCL T total	4.84 (0.79–8.89)	−2.79 (−6.78–1.19)	2.05 (−1.64–5.73)	0.09 (−0.30–1.15)
CBCL T internalizing ^d	6.36 (2.19–10.53)*	−4.76 (−8.74–0.77)*	1.60 (−2.11–5.32)	0.45 (−0.28–1.18)
CBCL T externalizing ^e	3.89 (−0.02–7.81)	−2.81 (−6.66–1.04)	1.08 (−2.46–4.62)	0.21 (−0.48–0.89)
TRF T total ^f	1.41 (−1.94–4.76)	−1.59 (−4.69–1.51)	−0.18 (−3.30–2.95)	−0.04 (−0.61–0.53)
TRF T internalizing ^g	2.22 (−1.74–6.17)	0.51 (−3.16–4.18)	2.73 (−0.97–6.42)	−0.05 (−0.63–0.72)
TRF T externalizing ^h	−1.58 (−4.52–1.36)	−0.50 (−3.26–2.27)	−2.08 (−4.85–0.70)	0.08 (−0.42–0.58)

* $P < 0.05$.^aCorrected for educational level of mother.^bCorrected for educational level of mother and gender.^cTotal number of atypical NEPSY scores. Corrected for gender, vanishing twin, Caesarian section and Apgar score 5 min.^dART group: no significant covariates, TTP: corrected for age of father.^eART groups: corrected for prematurity, foetal distress and age of mother. TTP: corrected for foetal distress and age of mother.^fCorrected for gender and vanishing twin.^gCorrected for vanishing twin.^hCorrected for vanishing twin and birthweight.

compared with the OS-IVF/ICSI and Sub-NC groups. In addition, we found that TTP was not associated with any of our cognitive and behavioural outcome parameters. These are reassuring results for both clinicians and parents involved in ART.

Our findings correspond with the majority of the literature (e.g. Cederblad et al., 1996; Wagenaar et al., 2008, 2009b; Bay et al., 2014; Fauser et al., 2014; Barbuscia and Mills, 2017), where no excess of abnormalities in cognitive development was found for children born after ART. We did find quite a high average IQ in our study group of 115 (SD 14.6), which is probably related to the relatively high educational level of the parents of our study group (41% of the parents had higher professional education or university education), which is often the case in this specific population of parents (Mac Dougall et al., 2012). Most studies, including our current results, also did not show clear consistent effects of ART on behavioural development (Cederblad et al. 1996; Wagenaar et al., 2008; Bay et al., 2014; Fauser et al., 2014; Barbuscia et al., 2019). These findings are in line with our earlier findings in the Groningen ART cohort at an age of 4 years (Schendelaar et al., 2016). An exception to this is the study by Wagenaar et al., (2009a, b) who found less externalising behaviour reported by the parents and more withdrawn and depressed behaviour rated by teachers in IVF children compared with a subfertile reference group. We did not find this, which could be due to the fact that in their study, the children were assessed at a later age (9–18 years) than in our study, and these types of behaviour might possibly become more pronounced in the teenage

years compared with the younger age-range. In contrast, we found lower CBCL internalising scores in the MNC-IVF/ICSI group compared with the OS-IVF/ICSI and Sub-NC groups. This probably reflects the special selection of parents who were able to complete the MNC-IVF/ICSI trajectory, as the MNC-trajectory requires long-lasting cooperation and perseverance of the parents (Pelinck et al., 2007). These parental characteristics may be associated with specific child-rearing attitudes and/or perception of child behaviour; perhaps the latter characteristic is most important as the advantage in internalising behaviour of the MNC-IVF/ICSI group was absent in the teachers' reports.

In our earlier studies in the Groningen ART cohort, we found that at the age of 4 years, the presence of subfertility, regardless of the application of ART, was associated with less optimal cognitive and behavioural outcomes (Schendelaar et al., 2016), although the absolute IQ differences and CBCL score differences between the study and fertile reference groups were small. A limitation of the current study is the lack of a fertile reference group to compare the cognitive and behavioural data of the subfertile group to. Nevertheless, the CBCL and TRF scores of our 9-year olds are in line with those of a non-referred normative sample of Achenbach and Rescoria (2001), indicating that the incidence of behavioural difficulties is at least within the normal range.

One of the main strengths of the Groningen ART cohort study is the study design, with two ART groups (OS-IVF/ICSI and MNC-IVF/ICSI) and a subfertile reference group, which allows for disentangling the

effects of OS and the *in vitro* laboratory procedures on developmental outcome. The current results indicate that both OS and the IVF laboratory procedures were not associated with worse cognitive or behavioural outcome at the age of 9 years. Another strength of this study is the long-term follow-up duration up to the age of 9 years, which is longer than in the majority of ART research (Rumbold *et al.*, 2017).

The attrition rate of 21% may be considered as a limitation of the study; however, after a follow-up period of 9 years, this rate is generally considered as acceptable (Fewtrell *et al.*, 2008). In addition, there were no significant differences in background characteristics between children who had and who did not have follow-up at the age of 9 years, making an attrition-related selection bias less likely.

Another limitation of the study is the relatively small sample size, which in itself also could contribute to selection bias and thereby hamper generalizability to the ART population. Also, it could lead to false negative findings as a result of underpowering. It would have been good to base the group sizes of our study on an *a priori* power analysis; however, as it was a longitudinal follow-up study, we had to deal with the groups formed on the basis of power calculations of the outcome at younger age. An *a priori* power analysis on total IQ indicated that both groups should contain 64 children (alpha 0.05, effect size 0.50, power 0.80), confirming that our study including 57 and 66 children in the OS-IVF/ICSI and Sub-NC groups, respectively, was slightly underpowered.

Concluding Remarks

This study did not find adverse effects of OS and the *in vitro* laboratory procedures on cognitive and behavioural outcome of children at the age of 9 years. In addition, TTP was not associated with cognition and behaviour at 9 years of age. These are reassuring results for the parents and clinicians involved in ART. Further studies should address follow-up beyond the age of 9 years. The period of puberty is important in brain development, as it involves great structural and functional changes, to prepare for adult-level functioning (Hadders-Algra, 2003; Herting *et al.*, 2017). Therefore, and notwithstanding the favourable findings of the present study, we strongly recommend extending the follow-up beyond this age into adolescence and adulthood.

Supplementary data

Supplementary data are available at *Human Reproduction* online.

Acknowledgements

We thank participating parents and children for their cooperation and enthusiasm; An Bennema, Pamela Schendelaar and Jorien Seggers for their assistance in the data collection; Anneke Kracht-Tilman and Linze Dijkstra for technical assistance.

Authors' roles

K.H. was involved in data-analysis and writing of the paper. D.K. was involved in data-collection and writing of the paper. S.L.B.G. was

involved in data-analysis and writing of the paper. M.J.H. was involved in study design and writing of the paper. M.H.-A. was involved in study design, data-collection, data-analysis and writing of the paper.

Funding

The study was financially supported by the UMCG, two graduate schools of the UMCG (BCN and SHARE) and the Cornelia Stichting. The sponsors of the study had no role in study design, data collection, data analysis, data interpretation or writing of the report.

Conflict of interest

The authors have no conflicts of interest to declare.

References

- Achenbach TM, Rescorla LA. *Manual for the ASEBA School-Age Forms and Profiles*. Burlington, VT: University of Vermont Research Center for Children, Youth, & Families, 2001
- Barbuscia A, Mills MC. Cognitive development in children up to age 11 years born after ART—a longitudinal cohort study. *Hum Reprod* 2017; **32**:1482–1488.
- Barbuscia A, Myrskylä M, Goisis A. The psychosocial health of children born after medically assisted reproduction: Evidence from the UK Millennium Cohort Study. *SSM Popul Health* 2019; **7**:100355.
- Bay B, Mortensen EL, Kesmodel US. Fertility treatment and child intelligence, attention, and executive functions in 5-year-old singletons: a cohort study. *BJOG* 2014; **121**:1642–1651.
- Carson C, Kelly Y, Kurinczuk JJ, Sacker A, Redshaw M, Quigley MA. Effect of pregnancy planning and fertility treatment on cognitive outcomes in children at ages 3 and 5: longitudinal cohort study. *BMJ* 2011; **343**:d4473.
- Carson C, Redshaw M, Sacker A, Kelly Y, Kurinczuk JJ, Quigley MA. Effects of pregnancy planning, fertility, and assisted reproductive treatment on child behavioral problems at 5 and 7 years: evidence from the Millennium Cohort Study. *Fertil Steril* 2013; **99**:456–63.
- Cederblad M, Friberg B, Ploman F, Sjöberg NO, Stjernqvist K, Zackrisson E. Intelligence and behaviour in children born after *in vitro* fertilization treatment. *Hum Reprod* 1996; **11**:2052–2057.
- Drenth Olivares M, Kuiper DB, Haadsma ML, Heineman KR, Heineman MJ, Hadders-Algra M. IVF procedures are not, but subfertility is associated with neurological condition of 9-year-old offspring. *Early Hum Dev* 2019; **129**:38–44.
- Fausser BC, Devroey P, Diedrich K, Balaban B, Bonduelle M, Delemarre-van de Waal HA, Estella C, Ezcurra D, Geraedts JP, Howles CM *et al.* Health outcomes of children born after IVF/ICSI: a review of current expert opinion and literature. *Reprod Biomed Online* 2014; **28**:162–182.
- Fewtrell MS, Kennedy K, Singhal A, Martin RM, Ness A, Hadders-Algra M, Koletzko B, Lucas A. How much loss to follow-up is acceptable in long-term randomised trials and prospective studies? *Arch Dis Child* 2008; **93**:458–461.
- Hadders-Algra M. Developmental coordination disorder: is clumsy motor behavior caused by a lesion of the brain at early age? *Neural Plast* 2003; **10**:39–50.

- Hart R, Norman RJ. The longer-term health outcomes for children born as a result of IVF treatment: part I—general health outcomes. *Hum Reprod Update* 2013a;**19**:232–243.
- Hart R, Norman RJ. The longer-term health outcomes for children born as a result of IVF treatment. Part II—mental health and development outcomes. *Hum Reprod Update* 2013b;**19**:244–250.
- Helmerhorst FM, Perquin DA, Donker D, Keirse MJ. Perinatal outcome of singletons and twins after assisted conception: a systematic review of controlled studies. *BMJ* 2004;**328**:261.
- Herting MM, Sowell ER. Puberty and structural brain development in humans. *Front Neuroendocrinol* 2017;**44**:122–137. <https://opendata.cbs.nl/statline/#/CBS/nl/https://www.nvog.nl/actueel/ivf-cijfers-2016-recordaantal-kinderen-geringe-toename-behandelcycli-en-meerlingen/>.
- Kapiteijn K, de Brijn CS, de Boer E, de Craen AJ, Burger CW, van Leeuwen FE, Helmerhorst FM. Does subfertility explain the risk of poor perinatal outcome after IVF and ovarian hyperstimulation? *Hum Reprod* 2006;**21**:3228–3234.
- Källén, Finnström OO, Lindam AP, Nilsson EM, Nygren KG, Otterblad Olausson PM. Is there an increased risk for drug treated attention deficit/hyperactivity disorder in children born after in vitro fertilization? *Eur J Paediatr Neurol* 2011;**15**:247–253.
- Knoester M, Helmerhorst FM, van der Westerlaken LA, Walther FJ, Veen S. Leiden artificial reproductive techniques follow-up project (L-art-FUP). Matched follow-up study of 5 8-year-old ICSI singletons: child behaviour, parenting stress and child (health-related) quality of life. *Hum Reprod* 2007;**22**:3098–3107.
- Korkman M, Kirk U, Kemp SL. Clinical NEPSY-II, Manual I. San Antonio, TX: Psychological Corporation, 2007a
- Korkman M, Kirk U, Kemp SL. NEPSY—Second Edition (NEPSY-II). San Antonio, TX: Harcourt Assessment, 2007b
- Mac Dougall K, Beyene Y, Nachtigall RD. 'Inconvenient biology:' advantages and disadvantages of first-time parenting after age 40 using in vitro fertilization. *Hum Reprod* 2012;**27**:1058–1065.
- Leunens L, Celestin-Westreich S, Bonduelle M, Liebaers I, Ponjaert-Kristoffersen I. Cognitive and motor development of 8-year-old children born after ICSI compared to spontaneously conceived children. *Hum Reprod* 2006;**21**:2922–2929.
- Leunens L, Celestin-Westreich S, Bonduelle M, Liebaers I, Ponjaert-Kristoffersen I. Follow-up of cognitive and motor development of 10-year-old singleton children born after ICSI compared with spontaneously conceived children. *Hum Reprod* 2008;**23**:105–111.
- Middelburg KJ, Haadsma ML, Heineman MJ, Bos AF, Hadders-Algra M. Ovarian hyperstimulation and the in vitro fertilization procedure do not influence early neuromotor development; a history of subfertility does. *Fertil Steril* 2010;**93**:544–553.
- Middelburg KJ, Heineman MJ, Bos AF, Hadders-Algra M. Neuromotor, cognitive, language and behavioural outcome in children born following IVF or ICSI—A systematic review. *Hum Reprod Update* 2008;**14**:219–231.
- Middelburg KJ, Heineman MJ, Bos AF, Pereboom M, Fidler V, Hadders-Algra M. The Groningen ART cohort study: ovarian hyperstimulation and the in vitro procedure do not affect neurological outcome in infancy. *Hum Reprod* 2009;**24**:3119–3126.
- Olivennes F, Rufat P, André B, Pourade A, Quiros MC, Frydman R. The increased risk of complication observed in singleton pregnancies resulting from in-vitro fertilization (IVF) does not seem to be related to the IVF method itself. *Hum Reprod* 1993;**8**:1297–1300.
- Pelincx MJ, Knol HM, Vogel NEA, Arts EGJM, Simons AHM, Heineman MJ, Hoek A. Cumulative pregnancy rates after sequential treatment with modified natural cycle IVF followed by IVF with controlled ovarian stimulation. *Hum Reprod* 2008;**23**:1808–1814.
- Pelincx MJ, Vogel NEA, Arts EGJM, Simons AHM, Heineman MJ, Hoek A. Cumulative pregnancy rates after a maximum of nine cycles of modified natural cycle IVF and analysis of patient drop-out: a cohort study. *Hum Reprod* 2007;**22**:2463–2470.
- Pinborg A. IVF/ICSI twin pregnancies: risks and prevention. *Hum Reprod Update* 2005;**11**:575–593.
- Rumbold AR, Moore VM, Whitrow MJ, Oswald TK, Moran LJ, Fernandez RC, Barnhart KT, Davies MJ. The impact of specific fertility treatments on cognitive development in childhood and adolescence: a systematic review. *Hum Reprod* 2017;**32**:1489–1507.
- Schendelaar P, Van den HER, Heineman MJ, La B-VGS, Middelburg KJ, Seggers J, Hadders-Algra M. Increased time to pregnancy is associated with less optimal neurological condition in 4-year-old singletons, in vitro fertilization itself is not. *Hum Reprod* 2014;**29**:2773–2786.
- Schendelaar P, La Bastide-Van Gemert S, Heineman MJ, Middelburg KJ, Seggers J, Van den Heuvel ER, Hadders-Algra M. Subfertility factors rather than assisted conception factors affect cognitive and behavioural development of 4-year-old singletons. *Reprod Biomed Online* 2016;**33**:752–762.
- Spangmose AL, Malchau SS, Schmidt L, Vassard D, Rasmussen S, Loft A, Forman J, Pinborg A. Academic performance in adolescents born after ART—a nationwide registry-based cohort study. *Hum Reprod* 2017;**32**:447–456.
- Wagenaar K, Huisman J, Cohen-Kettenis PT, Delemarre-van de Waal HA. An overview of studies on early development, cognition, and psychosocial well-being in children born after in vitro fertilization. *J Dev Behav Pediatr* 2008;**29**:219–230.
- Wagenaar K, van Weissenbruch MM, Knol DL, Cohen-Kettenis PT, Delemarre-van de Waal HA, Huisman J. Behavior and socioemotional functioning in 9–18-year-old children born after in vitro fertilization. *Fertil Steril* 2009a;**92**:1907–1914.
- Wagenaar K, van Weissenbruch MM, Knol DL, Cohen-Kettenis PT, Delemarre-van de Waal HA, Huisman J. Information processing, attention and visual-motor function of adolescents born after in vitro fertilization compared with spontaneous conception. *Hum Reprod* 2009b;**24**:913–921.
- Wagenaar K, van Weissenbruch MM, van Leeuwen FE, Cohen-Kettenis PT, Delemarre-van de Waal HA, Schats R, Huisman J. Self-reported behavioral and socioemotional functioning of 11- to 18-year-old adolescents conceived by in vitro fertilization. *Fertil Steril* 2011;**95**:611–616.
- Wechsler D. *Wechsler Abbreviated Scale of Intelligence WASI Manual*, 4th edn. Pearson/PsychCorp, San Antonio, USA. 1999.