Human Reproduction, Vol.32, No.2 pp. 447–456, 2017
Advanced Access publication on January 5, 2017 doi:10.1093/humrep/dew334

human reproduction

ORIGINAL ARTICLE ***Reproductive epidemiology***

Academic performance in adolescents born after ART—a nationwide registry-based cohort study

A.L. Spangmose[1,*], S.S. Malchau[1], L. Schmidt[2], D. Vassard[2], S. Rasmussen[1], A. Loft[3], J. Forman[4], and A. Pinborg[1]

[1]Department of Obstetrics and Gynaecology, Fertility Clinic, Copenhagen University Hospital, Hvidovre, Kettegaard Allé 30, 2650 Hvidovre, Copenhagen, Denmark [2]Section of Social Medicine, Department of Public Health, University of Copenhagen, Copenhagen, Denmark [3]Fertility Clinic Section 4071, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark [4]Section of Biostatistics, Department of Public Health, University of Copenhagen, Copenhagen, Denmark

*Correspondence address. Tel: +45 30 59 38 15; E-mail: annelrke@gmail.com

Submitted on November 21, 2016; resubmitted on November 21, 2016; accepted on December 7, 2016

STUDY QUESTION: Is academic performance in adolescents aged 15–16 years and conceived after ART, measured as test scores in ninth grade, comparable to that for spontaneously conceived (SC) adolescents?

SUMMARY ANSWER: ART singletons had a significantly lower mean test score in the adjusted analysis when compared with SC singletons, yet the differences were small and probably not of clinical relevance.

WHAT IS KNOWN ALREADY: Previous studies have shown similar intelligence quotient (IQ) levels in ART and SC children, but only a few have been on adolescents. Academic performance measured with standardized national tests has not previously been explored in a complete national cohort of adolescents conceived after ART.

STUDY DESIGN, SIZE, DURATION: A Danish national registry-based cohort including all 4766 ART adolescents ($n = 2836$ singletons and $n = 1930$ twins) born in 1995–1998 were compared with two SC control cohorts: a randomly selected singleton population ($n = 5660$) and all twins ($n = 7064$) born from 1995 to 1998 in Denmark. Nine children who died during the follow-up period were excluded from the study.

PARTICIPANTS/MATERIALS, SETTING, METHODS: Mean test scores on a 7-point-marking scale from −3 to 12 were compared, and adjustments were made for relevant reproductive and socio-demographic covariates including occupational and educational level of the parents.

MAIN RESULTS AND THE ROLE OF CHANCE: The crude mean test score was higher in both ART singletons and ART twins compared with SC adolescents. The crude mean differences were +0.41 (95% CI 0.30–0.53) and +0.45 (95% CI 0.28–0.62) between ART and SC singletons and between ART and SC twins, respectively. However, the adjusted mean overall test score was significantly lower for ART singletons compared with SC singletons (adjusted mean difference −0.15 (95% CI −0.29–(−0.02))). For comparison, the adjusted mean difference was +2.05 (95% CI 1.82–2.28) between the highest and the lowest parental educational level, suggesting that the effect of ART is weak compared with the conventional predictors. The adjusted analyses showed significantly lower mean test scores in mathematics and physics/chemistry for ART singletons compared with SC singletons. Comparing ART twins with SC twins yielded no difference in academic performance in the adjusted analyses. Similar crude and adjusted overall mean test scores were found when comparing ART singletons and ART twins.

LIMITATIONS, REASONS FOR CAUTION: Missing data on educational test scores occurred in 6.6% of adolescents aged 15–16 years for the birth cohorts 1995–1997, where all of the children according to their age should have passed the ninth grade exam at the time of data retrieval. As sensitivity analyses yielded no significant difference in the adjusted risk of having missing test scores between any of the groups, it is unlikely that this should bias our results. Adjustment for body mass index and smoking during pregnancy was not possible.

WIDER IMPLICATIONS OF THE FINDINGS: As our results are based on national data, our findings can be applied to other populations. The findings of this paper suggest that a possible small negative effect of parental subfertility or ART treatment is counterbalanced by the higher educational level in the ART parents.

STUDY FUNDING/COMPETING INTEREST(S): The Danish Medical Association in Copenhagen (KMS) funded this study with a scholarship grant. None of the authors had any competing interests.

TRIAL REGISTRATION NO. (STATISTICS DENMARK): 704676.

Key words: cognitive development / intelligence / academic performance / assisted reproductive technologies / IVF / ICSI / infertility / subfertility / ART

Introduction

Since the first baby was born after IVF in 1978, concern has been raised regarding the possible negative impacts of ART on the offspring.

Studies on long-term outcomes in ART children are inconsistent. While some studies have observed an increased risk of impaired behavioural or socio-emotional development, cognitive and psychomotor development and an increased risk of other mental disorders such as attention-deficit/hyperactivity disorder (ADHD) (Levy-Shiff *et al.*, 1998; Knoester *et al.*, 2008; Kallen *et al.*, 2011; Svahn *et al.*, 2015), other studies have found no increased risk of mental disorders, cognitive, neuromotor and emotional or behavioural development and no increased risk of neurological sequelae in ART children (Levy-Shiff *et al.*, 1998; Pinborg *et al.*, 2004; Ponjaert-Kristoffersen *et al.*, 2004, 2005; Zhu *et al.*, 2006; Ludwig *et al.*, 2009; Bay *et al.*, 2013). Cognitive development in ART children born after different conception methods, such as IVF and ICSI, has been explored with diverging conclusions, as some studies have found ICSI children to have delayed cognitive development (Bowen *et al.*, 1998; Knoester *et al.*, 2008; Goldbeck *et al.*, 2009) while others have shown no difference in IVF and ICSI children (Bonduelle *et al.*, 2003; Leslie *et al.*, 2003; Ponjaert-Kristoffersen *et al.*, 2005; Wennerholm *et al.*, 2006).

ART children have an increased risk of poor perinatal outcome such as preterm birth and low birthweight, mainly due to the high twin birth rates (Pinborg, 2005). But ART singletons also carry an increased risk of preterm birth and low birthweight as compared with their spontaneously conceived (SC) counterparts (Pandey *et al.*, 2012; Pinborg *et al.*, 2013; Ensing *et al.*, 2015). Moreover, ART treatment methods, such as oocyte donation and frozen embryo transfer, may result in altered perinatal outcomes (Malchau *et al.*, 2013; Pinborg *et al.*, 2013). The higher risk of poor perinatal outcome may affect the intelligence quotient (IQ) of the children negatively (Bhutta *et al.*, 2002; Shenkin *et al.*, 2004).

The Danish school system

All children entering public or private schools in Denmark are required to complete the same standardized written and oral exams after the ninth grade when finalizing lower secondary school.

The exam test scores are measured using a standardized 7-point-marking scale (from −3 to 12), that is directly comparable to the European Credit Transfer and Accumulation System (ECTS) grading scale (from F to A). ECTS is a marking scale made by the European Commission to facilitate comparability across borders. The standardized test scores are therefore comparable on a national and international level and an easier tool to measure academic performance in larger cohorts of adolescents compared with IQ tests. An average score is calculated for each student based on both the overall exam score and a teacher's score.

In Denmark, the ninth grade test scores give access to high schools which qualify for future higher level education; hence, the test scores are of great importance for the individual adolescent's educational opportunities. Since the exams are oral and written, academic as well as social skills influence the score level.

The aim of this study was to explore whether adolescents conceived after ART and aged 15–16 years achieve ninth grade test scores, which are comparable to test scores of SC adolescents and whether ART singletons and ART twins perform similarly.

Materials and Methods

Study population

This study compared a national register-based cohort, including an ART singleton and ART twin population, and two control populations, which have been described in more detail in previous publications (Fedder *et al.*, 2013; Malchau *et al.*, 2013). In short, the ART populations consist of all 4766 adolescents (n = 2836 singletons and n = 1930 twins) conceived after IVF or ICSI from 1 January 1995 to 31 December 1998. For the same study period, two control groups were selected from the Danish Medical Birth registry: (i) a random sample of SC singletons matched 2:1 on date of birth to the ART singleton population (n = 5660) and (ii) all SC twins born in the study period (n = 7064). We excluded adolescents conceived after frozen embryo transfer (n = 390) or oocyte donation (n = 80). As intrauterine insemination (IUI) treatments were not recorded in the IVF registry until 2007, adolescents conceived by IUI are included in the control groups. Our study population of adolescents born from 1995 to 1998 was selected to ensure that the great majority of the adolescents were aged 15–16 years and thereby had finished the ninth grade exam at the end of lower secondary school. We excluded nine children who had died between birth and follow-up (two ART singletons, one SC singleton and six SC twins).

National IVF registry

The Danish IVF registry is a mandatory cycle-based registry collecting information from all public and private fertility clinics in Denmark, including data on the type of ART treatment (IVF/ICSI) and treatment outcome. All citizens, who are born in Denmark, receive a personal identification number (PIN) immediately after birth, which is unique for that person. This PIN code is used in all national health registries. We used the maternal PIN code to cross-link data from the IVF registry with the Danish Medical Birth registry in terms of deliveries and outcomes. Data on perinatal outcomes were replaced by missing values if not included in the following intervals: gestational age 140–308 days, birthweight 200–6500 g and maternal parity <20.

According to Danish legislation, studies based solely on registry data, and with no personal involvement of the participants, do not require approval from a scientific ethics committee. The study was approved by the Danish Data Protection Agency (CVR no. 11-88-37-29 and Journal no. 2012-41-0848).

Outcomes and covariates

Based on maternal and child PIN codes, we retrieved data on academic test scores and socio-demographic characteristics from Statistics Denmark, which is the central authority of recording statistics on all Danish citizens.

Education of children up to adolescence is compulsory by law in Denmark. The public and private school systems have similar structure and include 10 years of education: 1 year of preschool and in total 9 years of primary and secondary school. After the ninth grade, all Danish adolescents aged 15–16 years are required to complete a general test of academic achievements, common for both public and private schools. The test covers mandatory subjects of academic achievements including Danish (oral, written), foreign languages (oral), mathematics (written) and physics/chemistry (oral). Scores for the general academic level throughout the year, a score given by the affiliated teachers, supplement each exam score.

The exam test and the score of the school year are marked by a 7-point-marking scale from −3 to 12. The average score is 7 and corresponds to the average score C of the ECTS scale. We used a mean test score based on a total mean of both the exam test scores and the teacher's scores of all the mandatory subjects and a specific mean test score for each of the mandatory subjects.

Because the overall average test score is obligatory and is used for admission to higher education in the Danish educational system, it is a valid measure to evaluate the general academic performance of adolescents.

Statistical analysis

Descriptive comparisons of ART and SC adolescents, singletons and twins, were made using Student's *t*-test for normally distributed continuous data, the Kruskal–Wallis test for non-normally distributed continuous data and the Chi-square test for categorical data.

Test scores were compared firstly in univariate analysis (Model 1) and secondly in a multivariate linear model (Model 2), which included the following covariates: maternal age (continuous variable), parity (<1 or ≥1), parents cohabiting at time of adolescents' graduation (yes/no), Danish ethnicity (yes/no), highest educational level of the parents (low, middle, high, highest), highest occupational level of the parents in five categories ((i) recipient of transfer income; (ii) other employee; (iii) employed, medium and lowest educational level; (iv) employed, highest educational level; (v) self-employed and chief executive), area of residence (five Danish regions), child sex (male/female) and graduation year (2010, 2011, 2012, 2013 and 2014). We also analysed the effect of parent educational level on the adjusted mean test score between the lowest and highest educational level. Birthweight and gestational age were not included in Model 2 since these factors are considered mediators on the causal pathway of ART and academic performance. To investigate whether a potential difference between ART and SC could be explained by the higher proportion of preterm births (gestational age <37 weeks) and low birth weights (birthweight <2500 g) in the ART cohort, these factors were added as explanatory variables together with the other confounders in Model 3.

Linear mixed models were used to account for the correlation in exam test scores of twins and siblings. The adequacy of the linear models was assessed by residual diagnostics.

IVF and ICSI adolescents, singletons and twins were compared in univariate (Model 1) and multivariate (Model 2) analysis. Sensitivity analyses were made for adolescents born in years 1995–1997 to ensure that results were not biased due to a higher proportion of missing test scores for adolescents born in 1998 as some of these children had not yet reached the age to pass the graduation tests at the time of data retrieval.

Finally, to investigate the risk of missing or failing (test score <2) the exams, we performed univariate and multivariate logistic regression analyses comparing the risk of a missing or failing test score in ART and SC adolescents. Multivariate analyses included the same confounding factors as the aforementioned except for 'graduation year'. We used generalized estimation equations to account for the correlation of missing and failed tests between twins and siblings.

Covariates were chosen *a priori* based on the current knowledge on factors influencing academic performance. We made a choice not to include family income as a covariate in the model due to substantial co-linearity with parental education and occupation.

All analyses were performed with IBM SPSS Statistics 19. A *P*-value <0.05 was considered statistically significant.

Results

Overall, 2836 ART singletons and 1930 ART twins were eligible for the study and 5660 SC singletons were included as controls. The 7064 SC twins born during the study period comprised the SC twin controls. For the birth cohorts 1995–1997, we obtained ninth grade test scores on 95.2% of the ART singletons, 92.6% of the SC singletons, 94.7% of the ART twins and 93.0% of the SC twins (Table I). For the children born in 1998, a lower proportion was registered with ninth grade test scores as some of the children would not have graduated at the time of data retrieval, because they were too young (Table I). In the crude analyses, significantly more ART singletons were registered with test scores compared with both SC singletons and ART twins (Table I).

However, after adjustment for confounding factors, the risk of not having test scores registered was equal in ART and SC singletons. ART twins were more likely to be recorded with missing test scores compared with SC twins with an adjusted odds ratio (aOR) 1.38 (95% CI 1.09–1.75). Comparing the risk of missing test scores in ART singletons vs ART twins showed an aOR of 0.76 (95% CI 0.61–0.95) (Table I). The sensitivity analyses including only adolescents born in 1995–1997 yielded no significant difference in the adjusted risk of having missing test scores between any of the groups.

Adolescent and parental background characteristics

As expected, compared with their SC counterparts, both ART singletons and twins were more likely to be born with low birthweight and preterm birth (Table II). ART adolescents were older at the time of graduation, they had significantly older mothers, they were more often firstborn and their parents were more often cohabiting at the time of the ninth grade graduation compared with SC adolescents (Table II). Further, parents of ART adolescents were more often of Danish ethnicity, and they had higher educational and occupational levels than parents of SC adolescents.

The ART singletons and ART twins differed with regard to low birthweight, preterm birth, age at the time of graduation, maternal age, parity and Danish ethnicity. Area of residence varied between ART singletons, ART twins, SC singletons and SC twins (Table II).

Test scores after the ninth grade

Crude overall mean test score for ART singletons was 7.16 (SD 2.41), which was significantly higher than for SC singletons 6.74 (SD 2.46), and the mean difference was +0.41 (95% CI 0.30–0.53) (Model 1) (Table III). Similarly, crude mean test score for ART twins was significantly higher than for SC twins 7.21 (SD 2.31) vs 6.78 (SD 2.50) (mean difference +0.45 (95% CI 0.28–0.62)) (Table III). No significant

Table I The proportion of the study population of children registered with test scores at school graduation.

	Singletons		Twins		Singletons ART vs SC		Twins ART vs SC		ART Singletons vs Twins	
	ART, *N* = 2836	SC, *N* = 5660	ART, *N* = 1930	SC, *N* = 7064	OR* (95% CI), *P*-value	aOR** (95% CI), *P*-value	OR* (95% CI), *P*-value	aOR** (95% CI), *P*-value	OR* (95% CI), *P*-value	aOR** (95% CI), *P*-value
1995, *N* (%)	495 (95.4)	941 (91.1)	366 (94.3)	1607 (92.4)						
1996, *N* (%)	632 (96.6)	1226 (93.0)	457 (96.0)	1682 (92.6)						
1997, *N* (%)	755 (94.0)	1503 (93.2)	450 (93.8)	1660 (94.0)						
Total *N* (%)	1882 (95.2)	3670 (92.6)	1273 (94.7)	4949 (93.0)	**0.62 (0.49; 0.79), *P* = 0.001**	0.98 (0.71; 1.34), *P* = 0.87	0.74 (0.54; 1.00), *P* = 0.05	1.12 (0.77; 1.62), *P* = 0.56	0.89 (0.63; 1.26), *P* = 0.51	0.83 (0.58; 1.20), *P* = 0.32
1998, *N* (%)	662 (77.0)	1315 (77.5)	405 (69.1)	1270 (72.9)						
Total *N* (%)	2544 (89.7)	4985 (88.1)	1678 (86.9)	6219 (88.0)	**0.85 (0.73; 0.98), *P* = 0.03**	1.17 (0.96; 1.42), *P* = 0.12	1.09 (0.90; 1.32), *P* = 0.39	**1.38 (1.09; 1.75), *P* = 0.01**	**0.77 (0.61; 0.96), *P* = 0.02**	**0.76 (0.61; 0.95), *P* = 0.02**

aOR, adjusted odds ratio; OR, odds ratio; SC, spontaneously conceived.
Statistically significant differences are shown in bold.
*Proportions of children not graduating the ninth grade exam were compared in logistic regression analyses using generalized estimation equations to adjust for correlation within twin pairs.
**Adjusted for the following confounders: maternal age, parity of the mother, cohabiting status, ethnicity, highest educational and occupational level of the parents, area of residence and child gender.

difference was observed between ART singletons and twins in the crude test scores (Table III).

In the multivariate regression analyses, ART singletons had significantly lower overall mean test scores (adjusted mean difference −0.15 (95% CI −0.29–(−0.02))) and significantly lower test scores in mathematics (adjusted mean difference −0.24 (95% CI −0.40–(−0.08))) and in physics/chemistry (adjusted mean difference −0.21 (95% CI −0.39 –(−0.03))) than SC singletons (Model 2) (Table III). The multivariate analyses did not yield any differences in test scores between ART and SC twins (Model 2 in Table III). ART singletons and twins had similarly adjusted mean test scores except for physics/chemistry, where ART singletons had a lower adjusted mean test score than ART twins (mean difference of −0.22 (95% CI −0.43–(−0.02))) (Model 2 in Table III).

IVF and ICSI adolescents stratified for singletons and twins had similar crude and adjusted mean test scores (Model 1) (Model 2) (Table SI).

In the pooled analyses of ART and SC singletons, the adjusted mean difference of test score was +2.05 (95% CI 1.82–2.28), higher for adolescents from families with the highest educational level vs families with the lowest educational level.

Birthweight and gestational age

Including preterm birth or low birthweight categories as covariates in the multivariate analyses did not make any significant changes in the mean difference of test scores between the groups compared to (Model 3) (data not shown).

Sensitivity analyses on adolescents born 1995–1997

Sensitivity analyses on mean test scores for adolescents born from 1995 to 1997 resulted in estimated adjusted mean differences that were similar to those based on the full cohorts. However, as the sample size was reduced, CIs became wider. In particular, the mean difference in overall test scores between ART and SC singletons was no longer statistically significant −0.13 (95% CI −0.28–0.03) similar to the difference in mean test score for physics/chemistry −0.21 (95% CI −0.42–0.00) (Table IV). For mathematics, the difference between ART and SC singletons remained significant −0.23 (95% CI −0.42–0.04).

In the sensitivity analyses, ART singletons achieved a lower mean test score in physics/chemistry than ART twins with an adjusted mean difference of −0.28 (95% CI −0.52–0.04) (Table IV).

Sensitivity analyses on adolescents with failing test scores (grade <2)

In the adjusted analysis, the proportions of adolescents with failing test scores (grade <2) were similar across the groups, except for ART singletons, who were more likely to fail an exam in mathematics, English and physics/chemistry compared with ART twins (Table SII).

Adolescents with total missing test scores

In all groups, adolescents with total missing test scores were less likely to have cohabiting parents and their parents had lower educational and occupational level than adolescents who were registered with test scores (Table SIII). Furthermore, they were more often born preterm or with low birth weight and were more often males.

SC singletons and ART twins with total missing test scores were more often of non-Danish ethnicity.

Discussion

This nationwide study is the first to compare exam test scores from a standardized national test achieved by adolescents conceived with or without ART treatment. The main finding was that ART adolescents achieved higher crude mean test scores than SC adolescents but after adjustment for parent's educational and occupational levels ART singletons had slightly lower overall mean test scores than SC singletons

Table II Socio-demographic characteristics and perinatal outcome of the study population and their parents reported at school graduation.

	Singletons ART, *N* = 2836	Singletons SC, *N* = 5660	Twins ART, *N* = 1930	Twins SC, *N* = 7064	*P*-value* Singletons ART vs SC	*P*-value* Twins ART vs SC	*P*-value* ART Singletons vs Twins
Child characteristics							
Gender (male/female), *N* (%)	1515/1321 (53.4)	2974/2686 (52.5)	1026/904 (53.2)	3658/3406 (51.8)	0.45	0.28	0.86
Birthweight, *N* (%)							
Low <2500 g	198 (7.0)	205 (3.6)	836 (43.3)	2825 (40.0)	**<0.001**	**0.008**	**<0.001**
Very low <1500 g	36 (1.3)	26 (0.5)	127 (6.6)	355 (5.0)	**<0.001**	**0.007**	**<0.001**
Gestational age, *N* (%)							
Preterm birth <37 weeks	150 (5.3)	181 (3.2)	495 (25.6)	1575 (22.3)	**<0.001**	**0.002**	**<0.001**
Very preterm birth <32 weeks	19 (0.7)	24 (0.4)	81 (4.2)	228 (3.2)	0.13	**0.04**	**<0.001**
Child age in year at test, mean (SD)	15.12 (0.40)	15.08 (0.45)	15.19 (0.43)	15.15 (0.44)	**<0.001**	**<0.001**	**<0.001**
Parents' characteristics							
Maternal age, years mean (SD)	48.8 (3.6)	44.5 (4.8)	48.5 (3.6)	45.5 (4.5)	**<0.001**	**<0.001**	**0.001**
Parity of the mother >1, *N* (%)	541 (19.1)	3250 (57.5)	415 (21.8)	4095 (58.8)	**<0.001**	**<0.001**	**0.03**
Cohabiting parents, *N* (%)	1956 (69.0)	3389 (59.9)	1311 (67.9)	4268 (60.4)	**<0.001**	**<0.001**	0.45
Danish ethnicity, *N* (%)	2773 (97.8)	5212 (92.1)	1907 (98.8)	6580 (93.1)	**<0.001**	**<0.001**	**0.009**
Highest educational level in the parents, *N* (%)							
Low	273 (10.0)	765 (14.2)	193 (10.4)	954 (14.1)	**<0.001**	**<0.001**	0.12
Middle	1119 (40.9)	2209 (41.1)	818 (44.0)	2848 (42.2)			
High	937 (34.3)	1754 (32.6)	584 (31.4)	2123 (31.5)			
Highest	406 (14.8)	648 (12.1)	264 (14.2)	822 (12.2)			
Highest occupational level of the parents, *N* (%)							
Employed, medium and lowest educational level	1287 (47.0)	2386 (44.3)	846 (45.5)	3156 (46.6)	**<0.001**	**<0.001**	0.49
Employed, highest educational level	509 (18.6)	924 (17.1)	378 (20.3)	1064 (15.7)			
Self-employed and chief executive	436 (15.9)	711 (13.2)	304 (16.3)	940 (13.9)			
Other employee	393 (14.4)	994 (18.4)	267 (14.4)	1138 (16.8)			
Recipient of transfer income	112 (4.1)	374 (6.9)	65 (3.5)	475 (7.0)			
Area of residence, *N* (%)							
North of Jutland	242 (8.5)	599 (10.6)	221 (11.5)	768 (10.9)	**<0.001**	**0.008**	**0.001**
Central Jutland	682 (24.0)	1346 (23.8)	491 (25.4)	1623 (23.0)			
Southern of Jutland and Fyn	616 (21.7)	1254 (22.2)	430 (22.3)	1574 (22.3)			
Copenhagen and North Zealand	907 (32.0)	1596 (28.2)	540 (28.0)	1973 (27.9)			
Central and South Zealand	389 (13.7)	865 (15.3)	248 (12.8)	1126 (15.9)			

*Data were compared with Student's *T*-test for normally distributed data, the Kruskal–Wallis test for non-normally distributed continuous data and the Chi-square test for categorical data.
Statistically significant differences are shown in bold.

Table III Crude and adjusted mean differences of test scores at school graduation for singletons and twins by mode of conception.

	Crude mean test score				Singletons ART vs SC		Twins ART vs SC		ART Singletons vs tTwins	
	Singletons		Twins		Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
	ART, N = 2544	**SC, N = 4985**	**ART, N = 1678**	**SC, N = 6219**	**Crude mean difference (95% CI), P-value**	**Adjusted mean difference (95% CI)*, P-value**	**Crude mean difference (95% CI), P-value**	**Adjusted mean difference (95% CI)*, P-value**	**Crude mean difference (95% CI), P-value**	**Adjusted mean difference (95% CI)*, P-value**
Overall	7.16 (2.41)	6.74 (2.46)	7.21 (2.31)	6.78 (2.50)	**0.41 (0.30; 0.53), P < 0.001**	**−0.15 (−0.29; −0.02), P = 0.03**	**0.45 (0.28; 0.62), P < 0.001**	−0.07 (−0.25; 0.10), P = 0.39	−0.03 (−0.19; 0.14), P = 0.75	−0.07 (−0.22; 0.08), P = 0.37
Danish	7.11 (2.44)	6.72 (2.47)	7.12 (2.35)	6.74 (2.48)	**0.39 (0.28; 0.51), P < 0.001**	−0.11 (−0.25; 0.02), P = 0.10	**0.39 (0.23; 0.56), P < 0.001**	−0.08 (−0.25; 0.09), P = 0.34	0.01 (−0.15; 0.18), P = 0.87	−0.01 (−0.17; 0.14), P = 0.86
Mathematics	7.19 (2.85)	6.79 (2.92)	7.28 (2.74)	6.82 (2.98)	**0.40 (0.26; 0.54), P < 0.001**	**−0.24 (−0.40; −0.08), P = 0.004**	**0.48 (0.28; 0.68), P < 0.001**	−0.09 (−0.29; 0.12), P = 0.42	−0.06 (−0.26; 0.13), P = 0.53	−0.12 (−0.31; 0.07), P = 0.21
English	7.60 (2.94)	7.12 (3.03)	7.50 (2.86)	7.05 (3.04)	**0.48 (0.34; 0.62), P < 0.001**	−0.12 (−0.29; 0.05), P = 0.17	**0.47 (0.26; 0.67), P < 0.001**	−0.12 (−0.33; 0.09), P = 0.28	0.11 (−0.09; 0.32), P = 0.27	0.05 (−0.15; 0.24), P = 0.64
Physics/ chemistry	6.85 (3.07)	6.48 (3.07)	7.06 (3.01)	6.70 (3.12)	**0.37 (0.22; 0.51), P < 0.001**	**−0.21 (−0.39; −0.03), P = 0.02**	**0.37 (0.16; 0.57), P < 0.001**	−0.14 (−0.36; 0.08), P = 0.22	−0.18 (−0.40; 0.03), P = 0.09	**−0.22 (−0.43; −0.02), P = 0.03**

Overall: mean test scores including all mandatory subjects (Danish, Mathematics, English, Physics/chemistry).
*Estimated differences in test scores obtained from a general linear model (singletons) and linear mixed model (twins) which included the following confounders:
Model 1: Unadjusted.
Model 2: Adjusted for the following confounders: maternal age, parity of the mother, cohabiting status, ethnicity, highest educational and occupational level of the parents, area of residence, child gender and graduation year.
Statistically significant differences are shown in bold.

because of lower mean scores in mathematics and physics/chemistry. Furthermore, academic performance in adolescence was similar in ART singletons and ART twins, in both the crude and adjusted analyses, except for the adjusted mean scores of physics/chemistry, where ART singletons performed slightly less well than ART twins.

IVF and ICSI adolescents had similar academic performance in both the crude and adjusted analyses for both singletons and twins.

Comparison with previous studies

An earlier study has shown a strong correlation between IQ and academic performance (Naglieri and Bornstein, 2003). In our study, we used academic performance as an outcome measure of cognitive development while most previous studies have examined IQ.

Several studies have assessed different aspects of cognitive development in ART preschool children up to the age of 5 years, while fewer studies have included school children and adolescents. The vast majority have adjusted for relevant confounders but have rather limited sample sizes and modest participation rates.

In preschool children, most studies have shown that IVF and ICSI children have psychomotor, cognitive or IQ level and behavioural development similar to that of spontaneously conceived (SC) children (Cederblad *et al.*, 1996; Leslie *et al.*, 2003; Place and Englert, 2003; Ludwig *et al.*, 2009; Sanchez-Albisua *et al.*, 2011; Bay *et al.*, 2014). An international collaborative study including children from several European countries is one of the largest on cognitive and motor development of 5-year-old ICSI-conceived children (n = 511) vs IVF (n = 424) and SC (n = 488) children (Ponjaert-Kristoffersen *et al.*, 2005). Apart from a few interaction effects between mode of conception and demographic variables, no differences were found in ICSI, IVF and SC children's scores in standardized tests of neuromotor and cognitive development. However, the authors conclude that demographic factors such as maternal educational level and maternal age may affect the cognitive development of IVF and ICSI children, compared with SC children (Ponjaert-Kristoffersen *et al.*, 2005).

For school children aged 8–10 years, results are contradictory which is explained by the major drawbacks of these studies, namely their limited sample sizes and participation rates and heterogeneous selection of controls leaving a high risk of bias. An Israeli study including term-born children at the age of 10 years concluded that IVF children (n = 51) have the same IQ as SC children (n = 51) (Levy-Shiff *et al.*, 1998). Two Belgian studies on ICSI children aged 8 years (n = 151) and 10 years (n = 109), respectively, compared with SC singletons (n = 109 aged 8 years and n = 90 aged 10 years) confirmed their results (Leunens *et al.*, 2006, 2008). One study from the Netherlands found lower IQ levels with children at the age of 5–8 years born after ICSI (n = 83) having lower scores than IVF children (n = 83) and an SC control population (n = 85) (Knoester *et al.*, 2008).

Only a few studies have examined academic achievements in adolescents. A Dutch cohort study (n = 233 IVF; n = 233 SC) and a US study (n = 308 IVF; n = 423 SC), including children and adolescents, showed good academic achievements in IVF children and adolescents compared with their SC counterparts (Wagenaar *et al.*, 2008; Mains *et al.*, 2010). In the Dutch study, including 8- to 18-year-old children, the participation rate was only 69% in IVF vs 51% in SC children (Wagenaar *et al.*, 2008). The US study showed that the IVF offspring aged 8–17 years scored higher than their matched peers in standardized tests across all

Table IV Adjusted mean differences of test scores at school graduation stratified on birth period for singletons and twins by mode of conception.

	Singletons ART vs SC				Twins ART vs SC				ART Singletons vs Twins			
	1995–1998		1995–1997		1995–1998		1995–1997		1995–1998		1995–1997	
	Adjusted mean difference (95% CI)*	*P*-value	Adjusted mean difference (95% CI)*	*P*-value	Adjusted mean difference (95% CI)*	*P*-value	Adjusted mean difference (95% CI)*	*P*-value	Adjusted mean difference (95% CI)*	*P*-value	Adjusted mean difference (95% CI)*	*P*-value
Overall	**−0.15 (−0.29; −0.02)**	**0.03**	−0.13 (−0.28; 0.03)	0.11	−0.07 (−0.25; 0.10)	0.39	−0.08 (−0.28; 0.11)	0.41	−0.07 (−0.22; 0.08)	0.37	−0.11 (−0.29; 0.07)	0.22
Danish	−0.11 (−0.25; 0.02)	0.10	−0.07 (−0.22; 0.08)	0.37	−0.08 (−0.25; 0.09)	0.34	−0.11 (−0.30:0.08)	0.26	−0.01 (−0.17; 0.14)	0.86	−0.03 (−0.21; 0.14)	0.70
Mathematics	**−0.24 (−0.40; −0.08)**	**0.004**	**−0.23 (−0.42; −0.04)**	**0.02**	−0.09 (−0.29; 0.12)	0.42	−0.05 (−0.29; 0.19)	0.70	−0.12 (−0.31; 0.07)	0.21	−0.20 (−0.42; 0.02)	0.07
English	−0.12 (−0.29; 0.05)	0.17	−0.08 (−0.29; 0.12)	0.42	−0.12 (−0.33; 0.09)	0.28	−0.14 (−0.38; 0.11)	0.28	0.05 (−0.15; 0.24)	0.64	0.04 (−0.19; 0.27)	0.72
Physics/ chemistry	**−0.21 (−0.39; −0.03)**	**0.02**	−0.21 (−0.42; −0.00)	0.05	−0.14 (−0.36; 0.08)	0.22	−0.15 (−0.40; 0.11)	0.26	**−0.22 (−0.43; −0.02)**	**0.03**	**−0.28 (−0.52; −0.04)**	**0.02**

Overall: mean test scores including all mandatory subjects (Danish, Mathematics, English, Physics/chemistry).
*Mean differences in test scores were compared in general linear model and linear mixed model with adjustment for the following confounders: maternal age, parity of the mother, cohabiting status, ethnicity, highest educational and occupational level of the parents, area of residence, child gender and graduation year.
Statistically significant differences are shown in bold.

grades and subtests, but a great drawback was a decline rate of 40.6% in their study (Mains *et al.*, 2010). In line with our study, they showed that parental educational level, maternal age and marital status were significant predictors of high academic performance. Furthermore, an Australian study based on structured telephone interviews with mothers and adolescents explored self-reported educational achievement comparing ART ($n = 705$) with SC ($n = 868$) adolescents yielded similar educational achievements in young adults aged 18–28 years (Halliday *et al.*, 2014). In the Australian study, 80% of the ART mothers were traced and of those only 55% participated.

A Dutch study including children aged 12 years ($n = 139$ IVF; $n = 143$ SC) measured information processing, attention and visual-motor function through various neuropsychological tests and then related these data to blood pressure and glucose levels (Wagenaar *et al.*, 2009). They showed no differences between IVF and control adolescents apart from a slight difference between the two groups for motor speed, although these scores were within the normal range for the test. No direct association was found between the cognitive measures and cardiometabolic outcomes (Wagenaar *et al.*, 2009). Response rates were similar to Wagenaar *et al.* (2008); 69% and 51% in IVF and controls, respectively.

Most of the aforementioned studies are included in a comprehensive systematic review by Haart and Norman (2013), which concluded that in general, the longer term mental and emotional health outcomes for IVF children are reassuring, and are very similar to those of SC children (Hart and Norman, 2013). This systematic review included 87 studies on mental health and development outcomes in ART offspring but very few studies on academic performance in adolescence. As the review includes all relevant literature, it is worthwhile mentioning that their results are in line with our results, showing few and small differences between ART and SC children in mental health overall.

Academic performance in twins

Two large cohort studies have shown that twins born before and in the 1950s have cognitive disadvantages compared with singletons (Deary *et al.*, 2005; Ronalds *et al.*, 2005). However, a Danish registry study explored ninth grade test scores in twins ($n = 3411$) and singletons ($n = 7796$) and found no significant differences except that twins performed slightly better in mathematics +0.13 (95% CI 0.03–0.23) (Christensen *et al.*, 2006), which is in line with our finding of better mean test scores in physics/chemistry in ART twins compared with ART singletons. They suggested that improvement of obstetric and paediatric practices may have ameliorated the earlier identified cognitive disadvantage in twins. However, Christensen *et al.* did not include information on conception method. In our study, the better test score in physics/chemistry in ART twins vs ART singletons may be explained by the higher mean age of twins at the time of graduation. In contrast, a Chinese study comparing preschool ART twins with ART singletons found significant lower cognitive developments among twins born after IVF but not after ICSI (Xing *et al.*, 2014). This may be due to the fact that younger children were included in the Chinese study, while the Danish studies looked at adolescent twins, who might catch up in academic performance over the years. All of the aforementioned studies made adjustments for various confounders including educational level of the parents.

Academic performance in IVF and ICSI adolescents

Previous studies regarding comparisons of IQ and academic performance of IVF and ICSI children aged 1–10 years are inconsistent. Three smaller studies with sample size <90 in each group found that ICSI children have a delayed cognitive function (Bowen *et al.*, 1998; Knoester *et al.*, 2008; Goldbeck *et al.*, 2009) compared with IVF children, while some larger studies showed no difference in cognitive function between IVF and ICSI children (Bonduelle *et al.*, 2003; Leslie *et al.*, 2003; Ponjaert-Kristoffersen *et al.*, 2005; Wennerholm *et al.*, 2006). The inconsistency between studies may be due to small sample sizes.

Our study is the largest published on academic performance in adolescents aged 15–16 years, and our results are in line with the previous largest studies which conclude that IVF and ICSI children have similar cognitive development.

Strengths and limitations

The strengths of our study are the national cohort design and the use of a standardized test performed by all students ending secondary school in Denmark, making this study the largest worldwide on academic performance in ART adolescents with a very high participation rate of 93.4% of the adolescents born between 1995 and 1997.

Additionally, we were able to take into account a wide range of relevant confounders for adolescents and parents, which is of importance as the study and control groups showed heterogeneity regarding most of these parameters. These differences in background data were expected, as earlier studies have found higher socioeconomic status and mean age in parents who conceived after ART treatment compared with the general population (Schmidt *et al.*, 1995).

A recent Danish cohort study exploring predictors of intelligence at the age of 5 years showed that parental education and maternal IQ are the most important predictors of IQ level of the child (Eriksen *et al.*, 2013), which confirms a previous German study (Shenkin *et al.*, 2004).

High maternal BMI and smoking during pregnancy are parameters that have shown a negative impact on long-term child outcomes such as IQ (Kupka *et al.*, 2014). As we were not able to adjust for maternal IQ, BMI and smoking, these parameters may cause residual confounding. However, BMI and smoking are correlated with socioeconomic factors which we do adjust for (Eek *et al.*, 2010; Toft *et al.*, 2015). The Danish follow-up study by Bay *et al.* (2014) exploring IQ in 5-year-old ART children found no differences in maternal IQ, self-reported maternal BMI or smoking between ART and SC mothers. Hence, we do not expect these parameters to be of considerable influence in our study.

Similar to previous studies, our results revealed that educational and occupational level of the parents has the highest impact on the academic performance of the offspring (Shenkin *et al.*, 2004; Eriksen *et al.*, 2013). For ART and SC singletons from families with the highest educational level, the adjusted mean difference in test score was +2.05 (95% CI 1.82–2.28) higher than in families with lowest educational level (data not shown). Additionally, a high occupational level of the parent, being the firstborn, having cohabiting parents, Danish ethnicity and being of female gender were also significant predictors of higher test scores in the ART vs SC singleton multivariate comparisons, more significant than ART treatment. Higher maternal age also had a significantly positive influence on academic performance in the children.

To avoid bias from other ART methods with different perinatal risk profiles, we excluded children conceived by oocyte donation and frozen embryo transfer (Malchau *et al.*, 2013; Pinborg *et al.*, 2013).

In our analyses, explained variance (R^2) was <1% in the univariate and 21% in the multivariate analyses, suggesting that the included covariates were important (data not shown).

Missing test scores

Of the adolescents born from 1995 to 1997, 93.4% were registered with educational test scores after the ninth grade, which is higher than the registration level of 84% in the previous Danish study including adolescents from 1986 to 1988 (Christensen *et al.*, 2006). Nevertheless, 6.6% of the adolescents born between 1995 and 1997 had missing data on test scores. The primary cause for the missing data on test scores is the varying quality of the school's level of registration of the test scores, although reporting test scores is statutory by law and no systematic missing reporting occurs. A few adolescents do not finalize the ninth grade, due to mental or physical disabilities but unfortunately, we have no information on this. Another group with no exam scores is adolescents who immigrate or pass the ninth grade abroad.

Only 74.8% of the adolescents born in the year 1998 had test scores recorded, which is explained by the fact that some of these adolescents were too young to have yet graduated when the data were collected.

Although the frequency of missing test scores was 6.6%, it should be emphasized that the risk of selection bias in our study is considerably lower than in smaller case–control studies with highly selected ART and control groups, which account for the great majority of the existing studies on academic performance.

ART twins born from 1995 to 1998 were less likely to be recorded with test scores compared with ART singletons and SC twins. Twins in general and especially ART twins were older when completing the ninth grade compared with singletons thus more of the twins born in 1997 and 1998 had not yet finished school at the time of data collection, which may be part of the explanation for the higher proportion of individuals with missing test scores among ART twins.

Sensitivity analyses on adolescents born in 1995–1997 showed no significant difference in the risk of missing test scores between ART singletons and twins, which supports the hypothesis that twins graduate from secondary school at a higher mean age and therefore fewer twins born in 1998 had passed at the time of data retrieval.

Another explanation could be linked to the poorer perinatal outcome of ART twins as more twins may have a mental or physical disability hindering them in finalizing secondary school.

The effect of prematurity on basic school achievement

Previous studies have found a negative association between preterm birth, low birthweight and later school performance (Shenkin *et al.*, 2004). This suggests that the negative effect of ART could be due to the higher proportion of children with poorer perinatal outcome in the ART cohort. However, including preterm birth and birthweight as covariates in our model did not change the estimated mean differences in test scores between ART and SC singletons, although the estimated effect of preterm birth and low birthweight confirmed the negative association. From this, we conclude that although preterm birth and

low birthweight has a negative impact on basic school performance, the number of premature ART children is too small to influence the academic performance of the ART population overall.

Conclusion

We found that ART singletons had slightly lower overall adjusted mean test scores than SC singletons. ART twins had similar adjusted scores to SC twins. ART singletons and twins performed similarly in both the crude and adjusted analyses.

Parental education was the factor with the highest impact on academic performance of the adolescents with an adjusted mean difference in test scores on 2.05 when comparing ART and SC singletons from families with the highest and lowest educational levels. Although we found significant differences in the adjusted mean test scores between ART and SC singletons, the differences were small and probably not of clinical relevance, well knowing we were not able to adjust for unknown variables.

Supplementary data

Supplementary data are available at *Human Reproduction* online.

Authors' roles

Substantial contribution was made by A.L.S. in terms of constructing the database, analysing and interpretation of the data and writing the first draft. S.S.M. contributed with statistical analyses and critical revision of the paper. L.S. contributed with the methodical design of the study and critical discussion and revision of the paper. D.V. contributed with statistical analyses and critical revision of the paper. S.R. contributed with the construction of the database and critical revision of the paper. A.L. contributed to critical discussion and revision of the paper. J.F. contributed with substantive statistical analyses and interpretation of the data and critical revision of the paper. A.P. contributed to the conception and design of the study and critically discussed and revised the paper. All authors finally approved the version being submitted.

Funding

The Danish Medical Association in Copenhagen (KMS) funded this study with a scholarship grant.

Conflict of interest

None declared.

References

Bay B, Mortensen E, Hvidtjorn D, Kesmodel U. Fertility treatment and risk of childhood and adolescent mental disorders: register based cohort study. *BMJ* 2013;**347**:f3978.

Bay B, Mortensen E, Kesmodel U. Fertility treatment and child intelligence, attention, and executive functions in 5-year-old singletons: a cohort study. *BJOG* 2014;**121**:1642–1651.

Bhutta AT, Cleves AM, Mario A, Casey PH, Cradock MM, Anand KJS. Cognitive and behavioral outcomes of school-aged children who were born preterm: a meta-analysis. *JAMA* 2002;**288**:728–737.

Bonduelle M, Ponjaert I, Van Steirteghem A, Derde MP, Devroey P, Liebaers I. Developmental outcome at 2 years of age for children born after ICSI compared with children born after IVF. *Hum Reprod* 2003;**18**:342–350.

Bowen JR, Gibson FL, Leslie GI, Saunders DM. Medical and developmental outcome at 1 year for children conceived by intracytoplasmic sperm injection. *Lancet* 1998;**351**:1529–1534.

Cederblad M, Friberg B, Ploman F, Sjoberg N, Stjernqvist K, Zackrisson E. Intelligence and behaviour in children born after in-vitro fertilization treatment. *Hum Reprod* 1996;**11**:2052–2057.

Christensen K, Petersen I, Skytthe A, Herskind A, McGue M, Bingley P. Comparison of academic performance of twins and singletons in adolescence: follow-up study. *BMJ* 2006;**333**:1095.

Deary I, Pattie A, Wilson V, Whalley L. The cognitive cost of being a twin: two whole-population surveys. *Twin Res Hum Genet* 2005;**8**:376–383.

Eek F, Ostergren P, Diderichsen F, Rasmussen N, Andersen I, Moussa K, Grahn M. Differences in socioeconomic and gender inequalities in tobacco smoking in Denmark and Sweden; a cross sectional comparison of the equity effect of different public health policies. *BMC Public Health* 2010;**10**:9.

Ensing S, Abu-Hanna A, Roseboom T, Repping S, van der Veen F, Mol B, Ravelli A. Risk of poor neonatal outcome at term after medically assisted reproduction: a propensity score-matched study. *Fertil Steril* 2015;**104**:384–390.

Eriksen H, Kesmodel U, Underbjerg M, Kilburn T, Bertrand J, Mortensen E. Predictors of intelligence at the age of 5: family, pregnancy and birth characteristics, postnatal influences, and postnatal growth. *PLoS One* 2013;**8**:e79200.

Fedder J, Loft A, Parner E, Rasmussen S, Pinborg A. Neonatal outcome and congenital malformations in children born after ICSI with testicular or epididymal sperm: a controlled national cohort study. *Hum Reprod* 2013;**28**:230–240.

Goldbeck L, Gagsteiger F, Mindermann I, Ströbele S, Izat Y. Cognitive development of singletons conceived by intracytoplasmic sperm injection or in vitro fertilization at age 5 and 10 years. *J Pediatr Psychol* 2009;**34**:774–781.

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* 2013;**19**:244–250.

Halliday J, Wilson C, Hammarberg K, Doyle L, Bruinsma F, McLachlan R, McBain J, Berg T, Fisher J, Amor D. Comparing indicators of health and development of singleton young adults conceived with and without assisted reproductive technology. *Fertil Steril* 2014;**101**:1055–1063.

Kallen A, Finnstrom O, Lindam A, Nilsson E, Nygren K, Otterblad Olausson P. Is there an increased risk for drug treated attention deficit/hyperactivity disorder in children born after in vitro fertilization? *EJPN* 2011;**15**:247–253.

Knoester M, Helmerhorst F, Vandenbroucke J, van der Westerlaken L, Walther F, Veen S. Cognitive development of singletons born after intracytoplasmic sperm injection compared with in vitro fertilization and natural conception. *Fertil Steril* 2008;**90**:289–296.

Kupka M, Ferraretti A, de Mouzon J, Erb K, D'Hooghe T, Castilla J, Calhaz-Jorge C, De Geyter C, Goossens V. Assisted reproductive technology in Europe, 2010: results generated from European registers by ESHREdagger. *Hum Reprod* 2014;**29**:2099–2113.

Leslie G, Gibson F, McMahon C, Cohen J, Saunders D, Tennant C. Children conceived using ICSI do not have an increased risk of delayed mental development at 5 years of age. *Hum Reprod* 2003;**18**:2067–2072.

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.

Levy-Shiff R, Vakil E, Dimitrovsky L, Abramovitz M, Shahar N, Har-Even D, Gross S, Lerman M, Levy I, Sirota L *et al*. Medical, cognitive, emotional, and behavioral outcomes in school-age children conceived by in-vitro fertilization. *JCCAP* 1998;**27**:320–329.

Ludwig A, Katalinic A, Thyen U, Sutcliffe A, Diedrich K, Ludwig M. Neuromotor development and mental health at 5.5 years of age of singletons born at term after intracytoplasmatic sperm injection ICSI: results of a prospective controlled single-blinded study in Germany. *Fertil Steril* 2009; **91**:125–132.

Mains L, Zimmerman M, Blaine J, Stegmann B, Sparks A, Ansley T, Van Voorhis B. Achievement test performance in children conceived by IVF. *Hum Reprod* 2010;**25**:2605–2611.

Malchau S, Loft A, Larsen E, Aaris Henningsen A, Rasmussen S, Andersen A, Pinborg A. Perinatal outcomes in 375 children born after oocyte donation: a Danish national cohort study. *Fertil Steril* 2013;**99**:1637–1643.

Naglieri J, Bornstein B. Intelligence and achievement: just how correlated are they? *JPA* 2003;**21**:244–260.

Pandey S, Shetty A, Hamilton M, Bhattacharya S, Maheshwari A. Obstetric and perinatal outcomes in singleton pregnancies resulting from IVF/ICSI: a systematic review and meta-analysis. *Hum Reprod Update* 2012;**18**: 485–503.

Pinborg A. IVF/ICSI twin pregnancies: risks and prevention. *Hum Reprod Update* 2005;**11**:575–593.

Pinborg A, Loft A, Schmidt L, Greisen G, Rasmussen S, Andersen A. Neurological sequelae in twins born after assisted conception: controlled national cohort study. *BMJ* 2004;**329**:311.

Pinborg A, Wennerholm U, Romundstad L, Loft A, Aittomaki K, Søderstrøm-Anttila V, Nygren K, Hazekamp J, Bergh C. Why do singletons conceived after assisted reproduction technology have adverse perinatal outcome? Systematic review and meta-analysis. *Hum Reprod Update* 2013;**19**:87–104.

Place I, Englert Y. A prospective longitudinal study of the physical, psychomotor, and intellectual development of singleton children up to 5 years who were conceived by intracytoplasmic sperm injection compared with children conceived spontaneously and by in vitro fertilization. *Fertil Steril* 2003;**80**:1388–1397.

Ponjaert-Kristoffersen I, Bonduelle M, Barnes J, Nekkebroeck J, Loft A, Wennerholm U, Tarlatzis B, Peters C, Hagberg B, Berner A *et al*. International collaborative study of intracytoplasmic sperm injection-conceived, in vitro fertilization-conceived, and naturally conceived 5-year-old child outcomes: cognitive and motor assessments. *Pediatrics* 2005;**115**:e283–e289.

Ponjaert-Kristoffersen I, Tjus T, Nekkebroeck J, Squires J, Verte D, Heimann M, Bonduelle M, Palermo G, Wennerholm UB, Collaborative study of Brussels, Göteborg and New York. Psychological follow-up study of 5-year-old ICSI children. *Hum Reprod* 2004;**19**:2791–2797.

Ronalds G, De Stavola B, Leon D. The cognitive cost of being a twin: evidence from comparisons within families in the Aberdeen children of the 1950 s cohort study. *BMJ* 2005;**331**:1306.

Sanchez-Albisua I, Lidzba K, Borell-Kost S, Mau-Holzmann U, Licht P, Krageloh-Mann I. Medical, psychological and intellectual development of 5-year-old children born after intracytoplasmic sperm injection. *Neuropediatrics* 2011;**42**:104–109.

Schmidt L, Munster K, Helm P. Infertility and the seeking of infertility treatment in a representative population. *BJOG* 1995;**102**:978–984.

Shenkin S, Starr J, Deary I. Birth weight and cognitive ability in childhood: a systematic review. *Psychol Bull* 2004;**130**:989–1013.

Svahn M, Hargreave M, Nielsen T, Plessen K, Jensen S, Kjaer S, Jensen A. Mental disorders in childhood and young adulthood among children born to women with fertility problems. *Hum Reprod* 2015;**30**:2129–2137.

Toft U, Vinding A, Larsen F, Hvidberg M, Robinson K, Glumer C. The development in body mass index, overweight and obesity in three regions in Denmark. *Eur J Public Health* 2015;**25**:273–278.

Wagenaar K, Ceelen M, van Weissenbruch M, Knol D, Delemarre-van de Waal H, Huisman J. School functioning in 8- to 18-year-old children born after in vitro fertilization. *Eur J Pediatr* 2008;**167**:1289–1295.

Wagenaar K, van Weissenbruch M, Knol D, Cohen-Kettenis P, Delemarre-van de Waal H, Huisman J. Information processing, attention and visual-motor function of adolescents born after in vitro fertilization compared with spontaneous conception. *Hum Reprod* 2009;**24**:913–921.

Wennerholm UB, Bonduelle M, Sutcliffe A, Bergh C, Niklasson A, Tarlatzis B, Kai CM, Peters C, Victorin Cederqvist A, Loft A. Paternal sperm concentration and growth and cognitive development in children born with a gestational age more than 32 weeks after assisted reproductive therapy. *Hum Reprod* 2006;**21**:1514–1520.

Xing L, Qian Y, Chen L, Zhang F, Xu X, Qu F, Zhu Y. Is there a difference in cognitive development between preschool singletons and twins born after intracytoplasmic sperm injection or in vitro fertilization? *J Zhejiang Univ Sci B* 2014;**15**:51–57.

Zhu J, Basso O, Obel C, Bille C, Olsen J. Infertility, infertility treatment, and congenital malformations: Danish national birth cohort. *BMJ* 2006; **333**:679.