Human Reproduction Vol.18, No.2 pp. 342–350, 2003 DOI: 10.1093/humrep/deg061

Developmental outcome at 2 years of age for children born after ICSI compared with children born after IVF

M.Bonduelle[1,5], I.Ponjaert[2], A.Van Steirteghem[3], M.-P.Derde[4], P.Devroey[3] and I.Liebaers[1]

[1]Centre for Medical Genetics, [2]Department of Developmental Psychology, [3]Centre for Reproductive Medicine and [4]Centre for Biostatistics, Dutch-speaking Brussels Free University, Brussels, Belgium

[5]To whom correspondence should be addressed at: Medische Genetica, AZ-VUB, Laarbeeklaan 101, B-1090 Brussels, Belgium. E-mail: maryse.bonduelle@az.vub.ac.be

BACKGROUND: Since the introduction of ICSI in 1991, medical outcome studies on ICSI children have been performed, but few have addressed developmental outcome. Hence, this outcome was assessed by performing a standard developmental test on children born after ICSI as compared with children born after IVF, at the age of 2 years. METHODS: In a prospective study, the medical and developmental outcome of 439 children born after ICSI (378 singletons, 61 twins) were compared with those of 207 children born after IVF (138 singletons, 69 twins), at the age of 24–28 months. These children were part of a cohort of children followed since birth. Of children reaching the age of 24–28 months between May 1995 and March 2002, 44.3% (2375/5356) were examined by a paediatrician who was unaware of the type of treatment used for each couple. Of all the children born, 12.2% (439/3618) in the ICSI group and 11.9% (207/1738) in the IVF group underwent a formal developmental assessment using the Bayley Scale of Infant Development (mental scale) by a paediatrician blinded to the type of treatment. RESULTS: There was no significant difference in maternal educational level, maternal age, gestational age, parity, birthweight, neonatal complication rate or malformation rate at 2 years between ICSI and IVF singletons, or between ICSI and IVF twins. No significant difference was observed in the developmental outcome using the Bayley scale at the age of 24–28 months (raw scores or test age) between ICSI children or IVF children. A multivariate regression analysis for the singleton children indicated that parity, sex (boys had lower scores than girls) and age had a significant influence on the test result, but that the fertility procedure (ICSI versus IVF) did not influence the test result. ICSI children from fathers with low sperm concentration, low sperm motility or poor morphology had a similar developmental outcome to that of children from fathers with normal sperm parameters. There were no significant differences between the initial cohort and the group lost to follow-up, nor between the psychologically tested and the non-tested group for a number of variables such as maternal educational level, birthweight in singletons and neonatal malformation rate. Although only some of the cohort of ICSI children were evaluated, a representative sample of both ICSI and IVF children was compared. CONCLUSIONS: There is no indication that ICSI children have a lower psychomotor development than IVF children. Paternal risk factors associated with male-factor infertility were found not to play a role in developmental outcome.

Keywords: children/development/ICSI/IVF/male and female infertility

Downloaded from https://academic.oup.com/humrep/article/18/2/342/639185 by guest on 20 August 2022

Introduction

ICSI was introduced at the authors' centre in 1991. As concerns were expressed about the safety of the new technique, a joint prospective study of the children born after ICSI was set up by the Centre of Reproductive Medicine and the Centre for Medical Genetics. For the IVF children conceived at the authors' centre, a prospective follow-up study of pregnancies and children had already been running since 1983. When ICSI was introduced in 1991 for couples with male-factor infertility, the follow-up for both types of assisted reproductive technology (ART) was conducted simultaneously. This led to the possibility of comparing the outcome of the ICSI procedure to the already well-established IVF procedure and so evaluating the risk related to the ICSI procedure (Bonduelle *et al*., 2002a). The aim of the present study was to collect data on karyotypes, congenital malformations, growth parameters, psychomotor development and fertility of the offspring, so as to evaluate possible adverse outcomes of the ICSI technique. The design of this follow-up study has been explained in several publications, and the study has led to satisfying follow-up rates and collaboration up to the age of 2 years, since parents were aware before planning a pregnancy of the emphasis placed on the follow-up studies (Bonduelle *et al*., 2002a). In order to quantify the developmental outcome at the age of 2 years (from

 © European Society of Human Reproduction and Embryology

24 to 28 months), the developmental outcome from May 1995 onwards was assessed by performing a standardized developmental test—the Bayley test—on children born after ICSI and IVF.

In the meantime, medical outcome studies on ICSI children have been published, but only a few have addressed the developmental outcome. Controversy arose with one publication in 1998, whereby a significantly lower score on the Bayley mental developmental index was reported for ICSI children at the age of 1 year (n = 89) than for a control group of children conceived by IVF (n = 84) or conceived naturally (n = 80) (Bowen *et al*., 1998). These data were not confirmed by an initial publication by the present authors' group on the basis of 221 ICSI children and 131 IVF children, where no difference was found between ICSI and IVF children on the Bayley scale (mental scale) at the age of 2 years (Bonduelle *et al*., 1998). Later publications (Sutcliffe *et al*., 1999, 2001) using the Griffiths' scale of mental development in a case–control study of 208 ICSI singletons compared with 221 singletons born after natural conception, did not show any significant difference between the ICSI children and their naturally conceived peers in terms of development at a mean age of 17 months.

In the present study, the data on medical and developmental outcome of ICSI children at their follow-up visit at the age of 2 years, as compared with a control group of IVF children followed in an identical prospective follow-up programme, were extended (Bonduelle *et al*., 1998, 2002a).

Materials and methods

Before starting a fertility treatment (IVF or ICSI), all couples were asked to agree to the follow-up conditions of the study, which include genetic counselling for the ICSI couples and participation in a prospective clinical follow-up study of the children. This study has been extensively described in previous publications (Bonduelle *et al*., 2002a,b). Indications for IVF were in particular tubal infertility, unexplained infertility, endometriosis and mild male-factor infertility. Indications for ICSI were oligoasthenoteratozoospermia, obstructive and non-obstructive azoospermia, and failed IVF (Devroey *et al*., 1995, 1996). Data on types of ART procedure and sperm parameters were recorded during the treatment phase. Data on pregnancies, deliveries and neonatal histories were retrieved from the written information obtained from gynaecologists, paediatricians and parents.

At follow-up examinations at 2 months, 12 months and 2 years, a detailed physical examination looking for major and minor malformations, together with a psychomotor evaluation, was performed by the same team of geneticist-paediatricians. Major and minor anomalies were registered using the ICD 10 code (Manual of the International Statistical Classification of Diseases, Injuries and Causes of Death, based on the 10th Revision) counting all malformations listed in the Q codes. Minor anomalies such as atrial septum defects (ASD), subluxation of the hips (Q65.5), pre-auricular tags (Q17.0) and torticollis (Q68.0) were also taken into account. Minor malformations (according to a checklist based on a previous report; Aase, 1990) without a Q code were omitted for this analysis of developmental outcome. One of the paediatricians, who had been trained at the Department of Clinical Psychology and was unaware of the type of ART used, tested the children on a Bayley Scale of Infant Development from which the mental scale was taken.

Between May 1995 and March 2002, 5356 ICSI and IVF children were conceived and born in the authors' hospital, from a total of 4391 pregnancies. Of the total of 5356 ICSI and IVF children, 2375 children attended the follow-up consultations at the age of 2 years. All of these children had a paediatric and psychomotor examination. For a limited cohort of 646 children, 439 ICSI and 207 IVF children reaching 2 years of age at the time of booking the appointment, a formal development assessment was made using the Bayley Scales of Infant Development from which a mental scale was tested, in order to quantify the psychomotor development of the children. This mental scale tests for perception, cognition and language, all three of which are part of early development in childhood. Motor items are tested in the motor scale, which was not tested here. The results were expressed as a difference between test ages and chronological ages and also calculated as K scores, with a score of <5 meaning a clinically significant developmental delay. In order to compare the results of ICSI and IVF children, different demographic and child variables (such as maternal age, maternal educational level, gestational age, parity, birthweight, neonatal complications and neonatal and actual malformations) were also registered (Table I)

The commonest reason for not participating in the Bayley evaluation was difficulty in fitting in with the limited time schedule of the paediatrician performing the developmental tests. In order to evaluate whether the two groups—children formally assessed by the Bayley scale versus children not assessed (but having a medical examination by the paediatrician)—were comparable, a number of variables that might possibly have influenced the developmental outcome in these groups were studied. These variables, such as maternal age, maternal educational level, gestational age, birthweight, neonatal complications and neonatal malformations, are detailed in Table II. Maternal educational level was split into six categories, the lowest level being 1. Levels 1, 2 and 3, from no education up to an educational level of 15 years, were grouped together as they were less frequent. Level 4 covered education up to the age of 18, i.e. completing secondary school; level 5 covered further (non-university) education, and level 6 included university degrees. For the children born from 2406 pregnancies, incomplete information was available up to 2 years. In this 'lost to follow-up group', the same variables (maternal age, maternal educational level, gestational age, birth order, Caesarean section, birthweight, neonatal malformations and neonatal complications) were also studied in comparison with the group which attended the follow-up consultations at 2 years, in order to evaluate whether the outcomes for ICSI and IVF children in the study group at 2 years was representative of the whole cohort of ICSI and IVF children conceived at the authors' centre (Table III).

Semen parameters of the fathers of ICSI children were evaluated according to procedures recommended by the World Health Organization, except for morphology, where strict Krüger criteria were used (World Health Organization, 1999). The cut-off values for abnormal sperm motility were <50% progressive motility (this is the total of a+b grade motility and less than the total motility, which is a+b+c). Semen samples with <14% normal morphology were considered abnormal. Different threshold values for sperm concentration ($\geq$20, <20, <15, <10, <5, <1 and <0.1$\times 10^6$/ml) were considered in the statistical analysis. Statistical analysis of the different sperm parameters in relation to Bayley scores in ICSI children was performed.

Statistical analysis

The statistical analysis was carried out using the SAS statistical program (SAS Institute Inc.). The testing was performed two-sided at the 5% level of significance.

Downloaded from https://academic.oup.com/humrep/article/18/2/342/639185 by guest on 20 August 2022

Table I. Comparison between ICSI and IVF children evaluated using the with Bayley scale of infant development between 24 and 28 months of age

	ICSI		IVF		Total	Statistical analysis
	Singletons	Twins	Singletons	Twins		
No. of children	439		207		646	
No. of singletons and twins (%)	378 (86.1)	61 (13.9)	138 (66.7)	69 (33.3)	646	
Maternal age at child birth (years)						
Mean ± SD	32.9 ± 4.2	31.6 ± 3.3	32.6 ± 3.6	31.1 ± 2.1		ANOVA Twins, $P = 0.001$; ICSI, $P = 0.275$ NS
Min–max	20.3–45.7	24.1–40.0	23.5–41.4	25.0–37.0		
Maternal educational level (weeks)[a]						
Σ lower educational levels (1, 2, 3, 4) (%)	41.7	37.3	50.0	46.1		ICSI, $P = 0.052$ NS
Σ higher educational levels (5, 6) (%)	58.3	62.7	50.0	53.9		
Duration of pregnancy (weeks)						
Mean ± SD	39.1 ± 1.5	35.1 ± 3.3	39.1 ± 1.5	35.5 ± 2.0		ANOVA: Twins, $P < 0.001$; ICSI, $P = 0.201$ NS
Min–max	31.6–41.5	27.2–39.4	32.2–41.5	31.2–39.5		
Parity						
Primiparity (%)	74.3	85.3	66.7	79.7		C-M-H: ICSI, $P = 0.058$ NS
Birthweight (g)						
Mean ± SD	3292 ± 512.3	2197 ± 631.4	3267 ± 462.6	2365 ± 387.3	3084 ± 643.4	ANOVA Twins, $P < 0.001$; ICSI, $P = 0.163$ NS
Min–max	1260–4610	900–3610	2270–4370	1415–3090	900–4610	
Sex						
Female (%)	52.1	50.8	50.7	47.8		C-M-H ICSI, $P = 0.681$ NS
n	197	31	70	33	331	
Male (%)	47.9	49.2	49.3	52.2		
n	181	30	68	36	315	
Neonatal complications						
%	20.9	78.7	18.8	79.7		C-M-H ICSI, $P = 0.708$ NS
n	79	48	26	55		
Malformation at 2 years ICD10 codes						
%	11.6	21.3	11.6	14.5		C-M-H ICSI, $P = 0.563$ NS
n	44	13	16	10	83	
% total	13.0		12.6			
n	57/439		26/207			
Bayley results						
Mean raw score ± SD	153.6 ± 6.07	149.2 ± 9.27	154.3 ± 5.75	149.6 ± 9.45		ANOVA Twins, $P < 0.001$; ICSI, $P = 0.425$ NS
K score <5 (%)	4.8		8.7			
K score <5 (*n*/total)	21/439		18/207			
K score <5	14 (3.7%)	7 (11.5%)	3 (2.2%)	15 (21.7%)		C-M-H Twins, $P < 0.001$; ICSI, $P = 0.525$
K score <5 corrected for term <36 weeks	14 (3.7%)	4 (6.6%)	3 (2.2%)	10 (14.5%)		C-M-H Twins, $P < 0.001$; ICSI, $P = 0.685$
Test age (months)						
Mean ± SD	27.3 ± 2.6	25.5 ± 3.5	27.5 ± 2.5	25.8 ± 3.7		ANOVA Twins, $P < 0.001$; ICSI, $P = 0.375$ NS
Min–max	17.5–30.0	16.0–30.0	20.5–30.0	18.0–30.0		
Age (months)						
Mean ± SD	24.8 ± 0.82	24.9 ± 0.98	24.8 ± 0.85	25.0 ± 0.93		ANOVA Twin, $P = 0.044$; ICSI, $P = 0.442$ NS
Min–max	24–28	24–28	24–28	24–28		
Difference in months	2.5	0.7	2.7	0.8		ANOVA Twins, $P < 0.001$; ICSI, $P = 0.520$ NS

[a]Maternal education: levels 1, 2, 3 = up to age 15 years; level 4 = up to 18 years; level 5 = further (non-university) education; level 6 = university degree.
$P < 0.05$ indicates statistically significant.
C-M-H = Cochran-Mantel-Haenszel test; NS = not significant.

Downloaded from https://academic.oup.com/humrep/article/18/2/342/639185 by guest on 20 August 2022

Table II. Analysis of drop-out at ⩾2 years: children seen at 2 years with Bayley records versus children without Bayley records

	Bayley test		Totals	Statistical analysis
	done	not done		
No. of:				
pregnancies	585	1416	2001	
Children	646	1729	2375	
Singletons	516	905	1421	
Twins	130	824	954	
Maternal age at child birth per mother (years)				
Mean ± SD	32.7 ± 4.01	32.9 ± 4.41		$P = 0.539$, Wilcoxon
Min–max	19.9–52.8	20.3–45.7		
Maternal educational level per mother *n* (%)				
Levels 1, 2, 3	16 (2.8)	18 (2.0)		
Level 4	236 (41.0)	433 (47.9)		
Level 5	245 (42.5)	308 (34.1)		
Level 6	79 (13.7)	145 (16.0)		
Total	576 (100.0)	904 (100.0)		$P = 0.222$, Wilcoxon
Duration of pregnancy/pregnancy (weeks)				
Mean ± SD	38.7 ± 2.05	37.7 ± 2.57		ANOVA: Twins, $P < 0.001$; Bayley, $P = 0.287$; Interaction, $P < 0.001$
Min–max	27.2–41.5	20.2–42.2		
Term in singletons (weeks)				
Mean ± SD	39.1 ± 1.50	38.7 ± 2.10		$P < 0.001$, *t*-test
Min–max	31.6–41.5	20.2–42.2		
Birthweight/pregnancy (g)				
Mean ± SD	3170.7 ± 588.1	2924.6 ± 682.4		ANOVA: Twins, $P < 0.001$: Bayley, $P = 0.314$; Interaction, $P = 0.001$
Min–max	900–4610	610–5280		
Birthweight in singletons (g)				
Mean ± SD	3285.4 ± 499.2	3217.9 ± 592.4		$P = 0.023$, *t*-test
Min–max	1260–4610	610–5280		
Neonatal complications/pregnancy *n* (%)	161 (27.5)	612 (43.3)		$P < 0.001$, Fisher's Exact test
Neonatal malformations ICD_{10}/pregnancy *n* (%)	81 (13.9)	149 (10.5)		$P = 0.038$, Fisher's Exact test
Age of children/pregnancy (years)				
Mean ± SD	2.07 ± 0.07	2.12 ± 0.12		$P < 0.001$, Wilcoxon
Range	2.0–2.4	2.0–2.4		
Sex of children *n* (%)				
Female singleton	267 (51.7)	458 (50.7)		C-M-H: Bayley done/not done, $P = 0.739$
Male singleton	249 (48.3)	446 (49.3)		
Female twin	64 (49.2)	405 (49.3)		
Male twin	66 (50.8)	417 (50.7)		
Procedure				
ICSI	439 (26.2)	1235 (73.8)		C-M-H: ICSI, $P = 0.030$; Twin, $P = 0.023$; Interaction, $P < 0.001$
IVF	207 (29.5)	494 (70.5)		
ICSI singleton	378 (37.0)	643 (63.0)		
IVF singleton	138 (34.5)	262 (65.5)		
ICSI twin	61 (9.3)	592 (90.7)		
IVF twin	69 (22.9)	232 (77.1)		

$P < 0.05$ indicates statistically significant.
C-M-H = Cochran-Mantel-Haenszel test.

The group of children seen only at birth was compared with the group of children who continued the follow-up programme up to the age of 2 years in terms of a number of characteristics. For the maternal characteristics, only one set of characteristics per pregnancy was considered in cases of twin pregnancies. These comparisons were performed using Fisher's Exact test for the binary variables and the Mann–Whitney test for the continuous and ordinal variables.

The group of children evaluated at 2 years (having a Bayley test or not) were also compared with the group not tested by a Bayley at 2 years for maternal and child characteristics. For the maternal characteristics, only one set of characteristics per pregnancy was considered in cases of twin pregnancies, and for neonatal complications and malformations, the incidence rate was also considered by pregnancy. The comparison of these variables was performed using Fisher's Exact test for the binary variables and the Mann–Whitney test for the continuous and ordinal variables. The child characteristics were compared using two-way ANOVA with Bayley testing (done/not done) and type of pregnancy (singletons/twins) as factors controlled. For the binary variables, use was made of the Cochran-Mantel-Haenszel (C-M-H) test controlled for type of pregnancy (singletons/twins). The Breslow–Day test was used in combination with the C-M-H test to assess the interaction between Bayley testing (done/not done) and type of pregnancy (singletons/twins). In case of a significant interaction the comparison of Bayley testing (done/not done) was performed separately for twins and singletons, using *t*-test and Fisher's Exact test respectively.

Downloaded from https://academic.oup.com/humrep/article/18/2/342/639185 by guest on 20 August 2022

Table III. Analysis of drop-out from birth: children with records at ⩾2 years versus children with no records at ⩾2 years

	Seen ⩾2 years	Not seen ⩾2 years	Total	Statistical analysis
No. of pregnancies	1985	2406	4391	
Maternal age at child birth				
No. observed	1980	2379	4359	
Mean (years)	32.8	33.1	33.0	$P = 0.024$, Mann–Whitney
Range (years)	19.9–52.8	20.6–52.0	19.9–52.8	
Maternal education level	*n* (%)	*n* (%)		
Levels 1, 2, 3	34 (2.3)	30 (1.7)		
level 4	661 (45.2)	796 (45.8)		
level 5	546 (37.3)	634 (36.4)		
level 6	223 (15.2)	280 (16.1)		
Total	1464 (100.0)	1740 (100.0)		$P = 0.697$, Wilcoxon
Duration of pregnancy				
No. observed	1898	2349	4247	
Mean (weeks)	38.0	37.7	37.8	$P < 0.001$ Wilcoxon
Parity (% primiparity)	72.8	71.6	72.2	$P = 0.377$, Fisher's Exact test
Caesarean section (%)	33.2	35.8	34.6	$P = 0.083$, Fisher's Exact test
Twin gestations (%)	28.6	34.6	31.9	$P < 0.001$, Fisher's Exact test
Birthweight/pregnancy				
No. observed	1977	2373	4350	
Mean ± SD (g)	3002 ± 665	2918 ± 690	2956 ± 680	$P < 0.001$, Wilcoxon
Range (g)	610–5280	660–5320	610–5320	
Neonatal complications/pregnancy (%)	36.8	38.2	37.6	$P = 0.362$, Fisher's Exact test
Neonatal malformation ICD_{10}/pregnancy (%)	10.1	6.1	7.9	$P < 0.001$, Fisher's Exact test
ICSI pregnancies (*n*)	1413	1619	3032	
IVF pregnancies (*n*)	572	787	1359	
ICSI pregnancies (%)	46.6	53.4	100	
IVF pregnancies (%)	42.1	57.9	100	
Cryopreserved pregnancies (%)	8.8	8.9	8.8	$P = 0.873$, Fisher's Exact test

$P < 0.05$ indicates statistically significant.

Finally, the type of treatment (ICSI/IVF) was compared for the group of children evaluated at 2 years having a Bayley test done, controlled for type of pregnancy (singleton/twin) using ANOVA for the continuous variables and the CMH test for the binary variables. Furthermore, a stepwise multivariate regression analysis was performed to predict the Bayley score in function of several potentially prognostic maternal and child characteristics in singleton children.

In the ICSI singleton children the relationship between sperm characteristics and the Bayley score was evaluated using the Mann–Whitney test and the Spearman's rank correlation coefficient.

Results

For the group of children lost to follow-up (seen only from birth to the age of 2 months) versus the children seen up to the age of 2 years, a number of maternal and infantile characteristics were compared (Table III). Maternal characteristics were compared per pregnancy. For 1985 pregnancies data were available on children up to the age of 2 years (follow-up rate of 45.2%); for 2406 pregnancies follow-up data (54.8%) on the children were not available up to the age of 2 years, but up to 2 months (in 58%) or to 1 year (Bonduelle *et al.*, 2002a). The follow-up rate up of children at 2 years was similar in the ICSI pregnancies (46.6%) and IVF pregnancies (42.1%). Maternal age was higher in the lost to follow-up group ($P = 0.024$). Maternal educational level, parity and percentage of neonatal complications were comparable in the two groups. Gestational age was lower in the lost to follow-up group ($P < 0.001$). The percentage of twin gestations was higher in the lost to follow-up group ($P < 0.001$) and percentages of Caesarean sections were similar in the two groups (P = NS). Birthweight was lower ($P < 0.001$) in the lost to follow-up group as well as the percentage ICD 10 codes and neonatal malformations ($P < 0.001$). The percentage of frozen and thawed embryos was the same in both groups: 8.8% of ICSI pregnancies and 8.9% of IVF pregnancies, for which follow-up data were available at 2 years of age.

For the groups of children formally assessed by the Bayley scale versus the children not assessed (but examined by the paediatrician), different variables possibly influencing the developmental outcome were compared (Table II). Maternal age was 32.7 years in the tested children versus 32.9 years in the untested children (P = NS). Maternal educational level was comparable in the two groups. Duration of pregnancy was 38.7 weeks in the tested children versus 37.7 weeks in the untested children (P = NS). Pregnancy duration in singletons was significantly different however ($P < 0.001$). Birthweight was 3170.7 g in the tested children versus 2924.6 g in the untested (P = NS) for the total group. In singletons, the birthweight was clinically similar, but slightly significantly different (P = 0.023) in both groups. Neonatal complications were recorded in 27.5% of the tested group versus 43.3% of the untested group (P < 0.001). Malformations coded in ICD 10 at birth were significantly more often present in the tested children (13.9%) than in the untested (10.5%) (P = 0.038). The sex ratio was the same in both groups. Of the ICSI children, 26.2% of the children (34.5% of the singletons) were seen at the follow-up consultations after 2 years, as were 29.5% of the IVF children (37.0% of the singletons). This percentage was lower in ICSI than in IVF (P = 0.030), but this was due to a higher percentage of twins seen in the IVF follow-up (twins $P = 0.023$ and interaction $P < 0.001$).

Downloaded from https://academic.oup.com/humrep/article/18/2/342/639185 by guest on 20 August 2022

Table IVA. Bayley results in relation to sperm morphology in ICSI children

	Sperm morphology		Statistical analysis
	⩾14%	<14%	
No. observed	71	177	Wilcoxon
Mean ± SD raw Bayley score	153.7 ± 5.4	153.4 ± 6.5	$P = 0.911$, NS
Min–max	136.0–162.0	124.0–162.0	

Table IVB. Bayley results in relation to sperm motility in ICSI children

	Motility		Statistical analysis
	⩾50%	<50%	
No. observed	107	202	Wilcoxon
Mean raw Bayley score	153.6	153.7	$P = 0.321$, NS
Min–max	137.0–162.0	124.0–163.0	

Table IVC. Bayley results in relation to sperm concentration in ICSI children

	Sperm concentration. ($\times 10^6$/ml)						
	<0.1	⩾0.1 to <1	⩾1 to <5	⩾5 to <10	⩾10 to <15	⩾15 to <20	⩾20
No. observed	11	20	63	35	25	21	149
Mean (± SD) raw Bayley score	152.5 ± 5.7	153.8± 6.9	153.9 ± 6.1	154.1 ± 5.2	153.8 ± 5.5	153.9± 6.6	153.1 ± 6.4
Min–max	141.0–159.0	136.0–162.0	136.0–162.0	144.0–163.0	138.0–161.0	138.0–161.0	124.0–162.0

Statistical analysis: Spearman rank correlation coefficient –0.05 no correlation.
NS = not significant.

As shown in Table I, 439 ICSI children (378 singletons and 61 twins) and 207 IVF children (138 singletons and 69 twins) underwent Bayley testing. For these children, maternal socio-demographic variables were similar for ICSI and IVF, controlled for singleton versus twin pregnancies, showing similar maternal ages, maternal educational levels and parities. Children conceived after ICSI versus IVF had similar pregnancy durations, birthweights, percentages of neonatal admissions or complications and malformation rates at 2 years. Similar sex ratios were apparent in both groups, and ages in both groups were the similar. Of the IVF children, 21 were born after replacement of a cryopreserved embryo, while of the ICSI children 40 were born after replacement of a cryopreserved embryo.

Between 24 and 28 months, a Bayley test was performed on 378 ICSI singletons and 61 ICSI twins, and on 138 IVF singletons and 69 IVF twins. Raw scores in ICSI and IVF singletons were similar, corresponding to a test age of 27.3 months in ICSI and 27.5 months in IVF, with a mean chronological age of 24.8 months (P = NS) in both groups. The test age was 2.5 months older than chronological age in ICSI singletons, and 2.7 months older in IVF singletons, which is a similarly elevated score (P = NS) as compared with the general population, and significantly different in singletons as compared with twins ($P < 0.001$). The ICSI twins scored slightly better (0.7 month) than the mean for the general population, as did the IVF twins (0.8 month), but there was no significant difference between ICSI and IVF. K scores of <5 were present for 21 ICSI children (4.8%) (14 singletons and seven twins) and for 18 IVF (8.7%) children (three singletons and 15 twins). After correction for prematurity under 36 weeks, 18 ICSI children (4.1%) (14 singletons and four twins) and 13 (6.3%) IVF children (three singletons and 10 twins) had K scores of <5. The difference in the percentage of K scores under 5 (in the singletons and twins) was statistically not significant. When corrected for prematurity, the rates remained similar for singletons but were lower for twins, confirming that IVF twins included the greatest number of children under the norm.

A multivariate regression analysis for the singleton children indicated that sex (boys had lower scores than girls) (P = 0.027), pregnancy duration (shorter pregnancy duration gave a lower score) (P = 0.001), parity (lower parity gave a better score) (P = 0.003) and child's age at testing (P = 0.002) had a significant influence on the test result, but that the procedure (ICSI versus IVF) did not influence the test result. Maternal educational level, birthweight, maternal age, malformations at birth, and neonatal complications were also taken into account, but did not contribute to the factors already retained in the model for prediction of the Bayley score.

The different threshold values for sperm concentration (⩾20, < 20, <15, <10, <5, <1 and <0.1 $\times$ 10^6/ml) were not associated with any difference in Bayley outcome (Table IV). Neither abnormal sperm motility nor sperm morphology

Downloaded from https://academic.oup.com/humrep/article/18/2/342/639185 by guest on 20 August 2022

showed a significant difference as regards the raw Bailey score in ICSI children.

Discussion

Counting from birth, a follow-up was obtained of 46.6% of ICSI children and 42.1% of IVF children at the 2-year examination. More children with a lower birthweight due to multiple pregnancies dropped out of the follow-up programme. Neonatal complications were similar, and neonatal malformations were even more frequent ($P < 0.001$) in the group continuing the follow-up (possibly due to a stricter registration of malformations for children attending the hospital). Maternal educational level was similar in both groups. A smaller proportion of twins was seen (28.6% of the initial cohort versus 34.6% not seen), but it is clear that there was no bias towards either ICSI or IVF singleton pregnancies in the follow-up (Table III). Twin children are often involved in other developmental follow-up programmes and are then less likely to continue the long-term ICSI and IVF follow-up programmes.

In the comparison of the groups of children formally assessed on the Bayley scale, as opposed to those children who had only a medical examination, different variables with a possible influence on the outcome show that important factors such as maternal educational level and birthweight (in singletons) were similar. Neonatal malformations were more frequent in the tested group ($P = 0.038$), which indicates that, far from selecting for a better group in the tested children, a group at greater risk seems to have been tested. Neonatal complications were more common in the untested group ($P < 0.001$) owing to a significantly higher percentage of twins (Table II). This also explains why both duration of pregnancy and consequently birthweight of the whole untested group, singletons and twins, were respectively shorter and lower. Pregnancy duration and birthweight for singletons, although being clinically similar, were statistically different between the tested and untested groups. A lower percentage of ICSI children than of IVF children were formally tested ($P = 0.030$), as a result of the lower percentage of ISCI twins completing the 2 years of follow-up ($P = 0.023$; interaction $P < 0.001$). From all of these data it can be deduced that a representative sample of ICSI and IVF singleton children was examined by means of a formal Bayley test. Only a small sample of the ICSI twins were formally tested.

As different percentages were found for the singleton and twin children, the singletons and twins were considered separately in Table I. All sociodemographic (maternal age, maternal educational level, parity) and child (duration of pregnancy, birthweight, neonatal complications, neonatal malformations) variables were similar (not significantly different) in the ICSI singleton group compared with the IVF singleton group, and in the ICSI and IVF twin groups. This means that comparable samples of the cohorts of ICSI and IVF children born in the authors' centre were indeed tested, with a bias towards more singletons seen at 2 years, and relatively more IVF twins than ICSI twins tested by means of a Bayley test. Of this initial cohort, it is also known from previous studies that the ICSI and IVF populations were similar (Bonduelle *et al.*, 2002a) in terms of parental criteria and child variables, such as percentage of multiple pregnancies, prematurity and birthweight. As all evaluated influencing factors, numbers of malformations at the time of examination and chronological ages were similar, it can be concluded that the nearly identical raw scores indicate that ICSI and IVF children perform at the same level on a standardized psychological test measuring perceptual, cognitive, verbal and conceptual early development. ICSI and IVF singletons perform comparably on the mental scale of the Bayley Scale of Infant Development, as do ICSI and IVF twins. The difference between singletons and twins however, was statistically relevant ($P < 0.001$).

These data were confirmed by the multivariate regression analysis (ANOVA) leading to a model for the singleton children indicating that sex (with girls scoring better), pregnancy duration (with shorter pregnancy duration leading to lower scores) and parity (lower parity gave a better score) were prognostic factors in the Bayley test result in ICSI as well as in IVF children, but these factors were already known from the literature to be predictors. ICSI and IVF procedures did not influence the test results. A number of parameters such as maternal educational level, maternal age, birthweight, malformations at birth, and neonatal complications did not contribute further to the model. Surprisingly, maternal educational level was not significantly related to the performance of the children, although the literature indicates that this is an important factor (Sternberg and Kaufman, 1998; Bornstein, 2002). A possible explanation might be that the maternal educational level in this study was more clustered in a homogeneous group (levels 4 and 5 accounting for 82.5%; see Table I) in the group followed for the full 2 years, tending to a higher level in both ICSI and IVF children, and that this may have obscured the effect of maternal educational level on the performance of the children.

A K score of <5 indicates the existence of clinically significant problems. In the singletons, there was no difference between ICSI (3.7%) and IVF (2.2%) children. IVF children in the present study scored similarly to the children tested in another study (Buitendijk, 2000) where 7.1% of the IVF singletons ($n = 197$) had a K score under 5 compared with 4.6% of the singletons in the general population ($n = 194$), which was not significantly different. If corrected for maternal age, numbers of older siblings, gestational age, sex of the child, maternal educational level, family situation and caring situation, the figures in the IVF and the control populations were identical. These very similar data in the Buitendijk study and the present study also indicate that there is probably no difference between ICSI and IVF and the general population. However, a direct study to compare children born after ICSI and in the general population, is needed to confirm this indirect conclusion.

Mean birthweights in children with K scores of <5 were 2794 g (median 2950 g) in the case of ICSI, and 2547 g (median 2500 g) in the case of IVF, which is lower than the average in the tested ICSI and IVF groups and may point to a relationship between lower performance and lower birthweight in these children.

Downloaded from https://academic.oup.com/humrep/article/18/2/342/639185 by guest on 20 August 2022

In the twins, the results for the number of children with K scores of <5 were lowest for the IVF twins, but this was not statistically significant; the mean test age for twins, however, was exactly the same for ICSI and IVF. Some 14% of IVF twins scored below the normal limit, and this might be related to the higher frequency of low birthweight observed in this group (Bonduelle *et al*., 2002a). In this limited twin group with K scores of <5, mean birthweight in twin children was 1807 g (median 1505 g) in ICSI and 2327 g (median 2300 g) in IVF. Of the 15 IVF twins, nine had a birthweight <2500 g. Here, birthweight might correlate with lower performance in this subgroup. These data are also in line with findings in the literature where IVF multiples (and singletons) have a higher risk of cerebral palsy and more developmental problems than the general population (Stromberg *et al*., 2002). None of the tested children in the present group had cerebral palsy, but since this study involves an incomplete cohort it cannot be stated with certainty that there were no children with cerebral palsy in the drop-out group.

The fact that twins perform worse than singletons has been well documented in the general population and is related to the higher prematurity rate, lower birthweight and possibly also the effect of 'ranking' within the family (Vandell *et al*., 1988). It is therefore clear that ICSI children taken as a whole will perform worse than the general population, owing to the higher rate of twin pregnancies. As long as ART produces more twins than in the general population, the developmental outcome of the children will be worse in ART. This is of course not a consequence of the ICSI procedure, or of the type of gametes used, but is simply a consequence of the policy in most ART programmes of transferring two or more embryos (Bonduelle *et al*., 2002a,b).

Males perform slightly but significantly worse than females ($P = 0.027$), with a mean raw Bayley score of 153.1 for ICSI and IVF boys, and 154.5 for ICSI and IVF girls. This slight difference might reflect the different rates of development between girls and boys at this age, for particular developmental items, as is known from different psychological studies (Nagy *et al*., 1999). Although this is generally accepted, the Bayley test does not differentiate for sex, probably because this is a somewhat older instrument or because it is not clinically relevant. It is not considered that this is a relevant fact specific to ICSI boys, as opposed to Bowen's conclusion that ICSI boys are more at risk of mild delays (Bowen *et al*., 1998), since in that study the difference between girls and boys was more obvious after ICSI than after IVF. In the present study, however, a similar lower performance by boys was observed after ICSI [154.0 $\pm$ 5.7 (SD) in girls, 153.2 $\pm$ 6.5 in boys] as after IVF (155.8 $\pm$ 5.6 in girls, 153.1 $\pm$ 5.6 in boys).

This study does not include testing of children from the general population. Test age is 2.5 months older than chronological age in ICSI singletons, and 2.7 months older in IVF singletons, which is the same increase in performance for the two groups as compared with the general population. However, this does not mean that ICSI children perform better than the general population. Indeed, there is a skewing due to the higher educational level of the parents undergoing fertility treatment (compared with the educational level in the general population) which biases the results towards higher scores. Moreover, the Bayley test was standardized some 20 years ago by testing a subset of 1283 children, tested for representativity of the Dutch population of children between 2 and 30 months on a sample of the general Dutch population (van der Meulen and Smrkovsky, 1983). It is known, however, that the average normal scores on intelligence tests and developmental tests have increased over time (Wolke *et al*., 1994). It can be concluded that on the one hand, ICSI and IVF children have a similar development, yet it is reported in the literature that IVF children are comparable with the general population (Buitendijk, 2000). To date, several studies have been published on psychomotor development with standardized tests at around 2 years of age in the IVF population as compared with the general population (Morin *et al*., 1989; Brandes *et al*., 1992; Raoul-Duval *et al*. 1994; Ron-El *et al.*, 1994; Cederblad *et al*., 1996; Buitendijk, 2000). Other studies have examined IVF children with no formal control group, comparing their results with the normal population as a reference (Mushin *et al*., 1986; Cederblad *et al*., 1996). One group (Morin *et al*., 1989) found a higher score for the IVF group on the Bayley test in a cohort of 83 IVF children matched to 93 non-IVF children, while others (Brandes *et al*., 1992) studied 116 IVF children and 116 matched controls by means of a Bayley test and a Stanford–Binet scale and found no difference in development between IVF children and the control group. Another group (Ron-El *et al*., 1994) compared 26 IVF children with a control group and found no difference using the General Cognitive Index test. Subsequently, a group of 197 IVF children was compared with 194 matched controls, and no difference was found on the Bayley scales (S.E.Buitendijk *et al*., unpublished data; doctoral thesis). It may be concluded indirectly therefore that there are no findings in the present study to indicate that ICSI children have a slower development than IVF children or, consequently, than the general population.

The present data are in line with the findings of other studies (Sutcliffe *et al*., 2001) which indicate that there is no significant difference between ICSI children and their naturally conceived peers in terms of development at a mean age of 17 months, using the Griffiths' scale of mental development in a case–-control study of 208 ICSI singletons compared with 221 singletons born after natural conception (Sutcliffe *et al*., 2001). However, the data are at variance with the results of another study (Bowen *et al*., 1998), in which the mean developmental index of the Bayley test was significantly lower for 89 ICSI children than for 84 IVF children and 80 children conceived naturally. In the latter study, however, singletons and twins were mixed (with a similar proportion in the three groups) and maternal educational level was significantly lower in the ICSI group, with more non-native-English-speaking mothers in the ICSI group. A subset analysis excluding the children from fathers with a lower unskilled profession was performed giving the same results. As it is assumed that maternal educational level is the most important predictive factor on infant development, it is considered that this analysis should have been performed in order to exclude this influence on the negative outcome of ICSI children. In the meantime, the group

Downloaded from https://academic.oup.com/humrep/article/18/2/342/639185 by guest on 20 August 2022

of children initially tested by Bowen and colleagues was followed until the age of 5 years by the same authors (Bowen *et al*., 1998; Leslie *et al*., 2002). Of the original cohort, 253 (84%) of the children were reassessed using the Wecshler Preschool and Primary Scale of Intelligence-Revised (WPPSI-R). For those children conceived after ICSI, all the IQ scales were within the normal range. There were no significant differences between the means for any IQ scales of the ICSI subgroup compared with either IVF or naturally conceived groups. ICSI did remain a weak predictor of full-scale IQ on univariate analysis, but this was not so once appropriate demographic co-variates were included in the analysis. These data are now reassuring and in line with the present findings, as well as with the data of others (Sutcliffe *et al*., 2001), indicating that a normal developmental outcome can be expected for ICSI children.

Analysis of different sperm parameters did not indicate any difference in the performance of ICSI children on the Bayley test. These data confirm the findings in the Sutcliffe study of 208 ICSI children, which found no significant differences in relation to the sperm characteristics of the fathers (Sutcliffe *et al*., 2001). It may be concluded therefore that there are no indications from the present study of 439 ICSI children, or from the literature data, that sperm quality influences the psychomotor development of ICSI children.

Acknowledgements

The authors are indebted to many colleagues: the clinical, scientific, nursing and technical staff of the Centre for Medical Genetics and the Centre for Reproductive Medicine, especially to the research nurses Andrea Buysse, Elke De Witte, Serena Debonnet, to Walter Meul and Hubert Joris for their efforts in compiling these data, and to Frank Winter of the Language Education Centre for reviewing the manuscript. These studies were supported by grants from the Fund for Scientific Research-Flanders, and an unconditional grant of Organon International; M.B. was also supported by the Scientific Fund Willy Gepts of the VUB University Hospital-AZ-VUB.

References

Aase, J.M. (1990) *Diagnostic Dysmorphology*. Plenum Medical Book Co., London, New York.

Bonduelle, M., Joris, H., Hofmans, K., Liebaers, I. and Van Steirteghem, A.C. (1998) Mental development of 201 ICSI children at 2 years of age. *Lancet*, **351**, 1553.

Bonduelle, M., Liebaers, I., Deketelaere, V., Derde, M.P., Camus, M., Devroey, P. and Van Steirteghem, A. (2002a) Neonatal data on a cohort of 2889 infants born after intracytoplasmic sperm injection (ICSI) (1991–1999) and of 2995 infants born after *in vitro* fertilization (IVF) (1983–1999). *Hum. Reprod.*, **17**, 671–694.

Bonduelle, M., Van Assche, E., Joris, H., Keymolen, K., Devroey, P., Van Steirteghem, A. and Liebaers, I. (2002b) Prenatal testing in ICSI pregnancies: incidence of chromosomal anomalies in 1586 karyotypes and relation to sperm parameters. *Hum. Reprod.*, **17**, 2600–2614.

Bornstein, M.H. (2002) Parenting infants. In Osofsky, J.D. and Fitzgerald, H.E. (eds), *WAIMH Handbook of Infant Mental Health*, Volume III. Wiley & Sons, New York, pp. 213–239.

Bowen, J.R., Gibson, F.L, Leslie, G.I. and Saunders, D.M. (1998) Medical and developmental outcome at 1 year for children conceived by intracytoplasmic sperm injection. *Lancet*, **351**, 1529–1534.

Brandes, J.M., Scher, A., Itzkovits, J., Thaler, I, Sarid, M. and Gershoni-Baruch R. (1992) Growth and development of children conceived by *in vitro* fertilization. *Pediatrics*, **90**, 424–429.

Buitendijk, S. (2000) *IVF pregnancies: outcome and follow-up*. Doctoral thesis, 7 June 2000, University of Leiden.

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

Devroey, P., Liu., J., Nagy, Z., Goossens, A., Tournaye, H., Camus, M., Van Steirteghem, A.C. and Silber, S. (1995) Pregnancies after testicular sperm extraction and ICSI in non-obstructive azoospermia. *Hum. Reprod.*, **10**, 1457–1460.

Devroey, P., Nagy, P., Tournaye, H., Liu, J., Silber, S. and Van Steirteghem, A.C. (1996) Outcome of intracytoplasmic sperm injection with testicular spermatozoa in obstructive and non-obstructive azoospermia. *Hum. Reprod.*, **11**, 1015–1018.

Leslie, G.I., Cohen, J., Gibson, F., McMahon, C. and Tennant, C. (2002) Children conceived using ICSI have normal development at school age. *Hum. Reprod.*, **18** (Abstract Bk.1), 3.

Manual of the International Statistical Classification of Diseases, Injuries and Causes of Death (ICD) (1992) *Based on the 10th Revision Conference*. World Health Organization, Geneva.

Morin, N.C., Wirth, F.H., Johnson, D.H., Frank, L. M., Presburg, H.J., Van de Water, V.L., Chee, E.M. and Mills, J.M. (1989) Congenital malformations and psychosocial development in children conceived by *in vitro* fertilization. *J. Pediatr.*, **115**, 222–227.

Mushin, D.N., Barreda-Hanson, M.C. and Spensley, J.C. (1986) *In vitro* fertilisation children: early psycho-social development. *J. In-Vitro Fertil. Embryo Transfer*, **3**, 247–252.

Nagy, P., Jacklin, C. and Martin, L. (1999) Effects of gender on behaviour and development. In Levine, M.D., Carey, W.B. and Crocker, A.C. (eds), *Developmental Behavioral Pediatrics*. W.B.Saunders Co., Philadelphia, pp. 100–106.

Raoul-Duval, A., Bertrand-Servais, M., Letur-Konirsch, H. and Frydman, R. (1994) Psychological follow-up of children born after in-vitro fertilization. *Hum. Reprod.*, **9**, 1097–1101.

Ron-El, R., Lahat, E., Golan, A., Lerman, M., Bukovsky, I. and Herman, A. (1994) Development of children born after ovarian superovulation induced by long-acting gonadotropin-releasing hormone agonist and menotropins, and by *in vitro* fertilization. *J. Pediatr.*, **125**, 734–737.

Stromberg, B., Dahlquist, G., Ericson, A., Finnstrom, O., Koster, M. and Stjernqvist, K. (2002) Neurological sequelae in children born after in-vitro fertilisation: a population-based study. *Lancet*, **359**, 461–465.

SAS Institute Inc. (1999) *SAS/STAT User's Guide*, SAS Institute, Version 8. Cary, NC, SAS Institute, Inc.

Sternberg, R.J. and Kaufman, J.C. (1998) Human abilities. *Annu. Rev. Psychol.*, **49**, 479–502.

Sutcliffe, A.G., Taylor, B., Li, J., Thornton, S., Grudzinskas, J.G. and Lieberman, B.A. (1999) Children born after intracytoplasmic sperm injection: population control study. *Br. Med. J.*, **318**, 704–705.

Sutcliffe, A.G., Saunders, K., Thornton, S., Lieberman, B.A. and Grudzinskas, J.G. (2001) Outcome in the second year of life after in-vitro fertilisation by intracytoplasmic sperm injection: a UK case-control study. *Lancet*, **357**, 2080–2084.

Vandell, D., Owen, M., Wilson, K. and Henderson, V. (1988) ''Social development in infant twins: peer and mother child relationships. *Child Dev.*, **59**, 168–177.

Van der Meulen, B.F. and Smrkovsky, M. (1983) *Bayley ontwikkelingsschalen*. Thesis, State University, Groningen.

World Health Organization (1999) *Laboratory Manual for the Examination of Human Semen and Sperm Cervical Mucus Interaction*. 4th edn, Cambridge University Press, Cambridge.

Wolke, D., Ratschinski, G., Ohrt, B. and Riegel, K. (1994) The cognitive outcome of very preterm infants may be poorer than often reported: an empirical investigation of how methodological issues make a big difference. *Eur. J. Pediatr.*, **153**, 906–915.

Submitted on August 8, 2002; accepted on October 17, 2002

Downloaded from https://academic.oup.com/humrep/article/18/2/342/639185 by guest on 20 August 2022