

Cognitive and psychomotor development of 5- to 6-year-old singletons born after PGD: a prospective case–controlled matched study

C. Winter^{1,2,*}, F. Van Acker³, M. Bonduelle², S. Desmyttere²,
F. De Schrijver², and J. Nekkebroeck^{1,2}

¹Department of Developmental and Lifespan Psychology, Vrije Universiteit Brussel (VUB), Pleinlaan 2, Brussels 1050, Belgium ²Centre for Medical Genetics, UZ Brussel, Vrije Universiteit Brussel (VUB), Laarbeeklaan 101, 1090 Brussels, Belgium ³Open Universiteit, Heerlen, The Netherlands

*Correspondence address. christiane.winter@uzbrussel.be

Submitted on January 3, 2014; resubmitted on May 15, 2014; accepted on June 4, 2014

STUDY QUESTION: Do preschool preimplantation genetic diagnosis (PGD) children differ in their cognitive and psychomotor development from children born after ICSI and spontaneous conception (SC)?

SUMMARY ANSWER: The cognitive development of PGD pre-schoolers was comparable to children born after ICSI and SC but motor development differed between ICSI and SC groups.

STUDY DESIGN, SIZE DURATION: The cognitive abilities and motor skills of 5- to 6-year-old singletons born after PGD ($n = 47$) were assessed in comparison with 49 ICSI and 48 SC children in a prospective, case–controlled, matched follow-up study between April 2011 and May 2013.

PARTICIPANTS/MATERIALS, SETTING, METHODS: PGD singletons, ICSI and SC children of preschool age were examined with the Wechsler Preschool and Primary Scale of Intelligence (WPPSI-III-NL) and the Movement ABC (M ABC). The WPPSI-III-NL revealed scores for Full IQ, Verbal IQ and Performance IQ. The M ABC yields a total score and comprising scores for measurements of balance, dexterity and ball skills. Since embryo biopsy is the only technical difference between the PGD and ICSI procedures, ICSI children were included as controls. These children were part of a Dutch-speaking cohort of children conceived after assisted reproduction technology (ART) at the Universitair Ziekenhuis Brussel (UZ Brussel) who received longitudinal follow-up. The SC children acted as a second control group similar to the fertile PGD sample and in contrast to the ICSI group. The SC group was recruited through announcements in a variety of media. The children were matched for age, gender, birth order and educational level of the mother. The assessments carried out for the ART groups were blinded whenever possible. The data were analysed using analysis of covariance (ANCOVA) and partial eta squared (η^2), which was used as a measurement of effect size.

MAIN RESULTS AND THE ROLE OF CHANCE: The overall cognitive development of PGD singletons did not differ from controls [$P = 0.647$, $\eta^2 = 0.006$; 95% confidence interval (CI) (0, 0.043)]. The partial IQ scores for Verbal and Performance intelligence revealed similar results. Analysis of motor development based on the total score as well as subscales did indicate a significant difference between the three conception groups [$P = 0.033$, $\eta^2 = 0.050$, 95% CI (0, 0.124)]. *Post hoc* analysis indicated that the significant difference was situated between performances of ICSI and SC children. Balance capacities [$P = 0.004$, $\eta^2 = 0.079$, 95% CI (0.025, 0.163)] and its *post hoc* analysis yielded equivalent results. Motor capacities of PGD singletons, however, did not differ from any of the two other conception groups.

LIMITATIONS, REASONS FOR CAUTION: Given that we only assessed Caucasian singletons born after PGD, caution is required when drawing more general inferences from our results. The small sample size may be a limitation. *A priori* power analysis, however, revealed that at least 52 children per group were needed to detect a medium effect and 80% power using ANCOVA. Originally our sample met this threshold but we had to exclude six cases in order to remove outliers and due to missing data.

WIDER IMPLICATIONS OF THE FINDINGS: Long-term follow-up of children born after embryo biopsy, in this case for PGD, is needed to confirm that the development of these children remains comparable to ICSI and SC children. Our findings do support the safety of the PGD technique and will reassure patients with hereditary genetic diseases regarding the health of their future offspring conceived with PGD.

STUDY FUNDING/COMPETING INTEREST(S): Funding for this study was obtained from the OZR (Research group of the Vrije Universiteit Brussel), the FWO (Fonds Wetenschappelijk Onderzoek) and the Wetenschappelijk Fonds Willy Gepts. The UZ Brussel and the Centre of Medical Genetics received funding from pharmaceutical firms for data collection. UZ Brussel and the Centre for Medical Genetics have received many educational grants for organizing the data collection, from IBSA, Ferring, Organon, Shering-Plough, Merck and Merck Belgium. M.B. has received consultancy and speaker's fees from Organon, Serono Symposia and Merck.

Key words: PGD / cognitive development / motor outcome / children / follow-up

Introduction

PGD is an established technique which has been performed for the last 20 years (Handyside, 2010) for parents at risk of transmitting monogenetic and chromosomal genetic disorders who want to avoid termination of pregnancy after prenatal diagnosis (Verlinsky *et al.*, 2004; Verpoest *et al.*, 2009). The PGD technique includes an ART. Embryos are obtained after IVF with ICSI (Basille *et al.*, 2009). In PGD, ICSI is used to enhance the likelihood of fertilization by injecting one sperm cell directly into one oocyte. The next phase includes a biopsy, which is performed on one or two blastomeres of the 6–8-cell embryo. The PCR or fluorescence *in situ* hybridization technique is used for diagnosis (Sermon *et al.*, 2004; Staessen *et al.*, 2004; Harper *et al.*, 2012). Unaffected embryos are then transferred into the uterus on Day 5 after oocyte retrieval.

Traditionally, follow-up studies of children born after ART have focused primarily on IVF/ICSI children. Systematic reviews of outcomes, analysing data for neuromotor and cognitive development as well as for language and behavioural disorders of children born after ART, show reassuring results (Izat and Goldbeck, 2008; Middelburg *et al.*, 2011). Authors (Bouwen *et al.*, 1998; Sutcliffe *et al.*, 1999, 2001; Place and Englert, 2003; Ponjaert *et al.*, 2005; Knoester *et al.*, 2007, 2008; Ludwig *et al.*, 2009) differentiated clearly between children born after IVF as opposed to ICSI children born after the more invasive ICSI technique. Some authors also considered the type of infertility and the technique used (female infertility: IVF versus male infertility: ICSI) which are hypothesized to play different roles in the very early development of human beings (Sutcliffe and Ludwig, 2007) and subsequently later in life (Fisher-Jeffes *et al.*, 2006; Middelburg *et al.*, 2008).

In a systematic review by Middelburg *et al.* (2008) on neuromotor and cognitive differences between children born after ICSI and spontaneous conception (SC), two studies are mentioned which showed lower intelligence quotient (IQ) scores after ICSI conception as opposed to SC (Bouwen *et al.*, 1998; Knoester *et al.*, 2008).

In contrast with IVF/ICSI related studies, the literature on exclusively PGD follow-up is scarce. The neonatal and medical outcomes of children born after PGD when compared with ICSI are encouraging (Desmyttere *et al.*, 2009; 2012; Liebaers *et al.*, 2010). To our knowledge, few studies have been published on the cognitive and motor development of children born after PGD and preimplantation genetic screening (PGS).

Research into the cognitive development of children born after PGD/PGS: has been reassuring although results were not always unanimous (Banerjee *et al.*, 2008, Nekkebroeck *et al.*, 2008). Banerjee *et al.*'s (2008) locomotor scores turned out to be significantly lower for PGD children than those for controls when compared with the findings of Nekkebroeck *et al.* (2008). Moreover, these two pioneering studies did not differentiate between PGD and PGS.

It has been argued that PGD and PGS children should be considered as different conception groups because PGD is used for genetic problems, whereas PGS is used for advanced maternal age, chromosomal and fertility problems (Nekkebroeck *et al.*, 2008). Recently, Dutch study groups have focused on samples that consisted exclusively of PGS children (Middelburg *et al.*, 2011; Schendelaar *et al.*, 2013) and found their cognitive development and neurodevelopment comparable to IVF children without PGS.

To our knowledge, only one recent study has focused solely on PGD children (Thomaidis *et al.*, 2012). Thirty-one children of different ages including singletons, twins and triplets were examined. Of these, 77.4% showed normal general development, while 22.6% of the children showed mild motor retardation quotients. One major limitation of this assessment was the lack of matched control groups for the developmental examinations.

Methodologically, correct follow-up research with matched control groups of comparable age in children born after PGD is scarce. Long-term follow-up research in children born after PGD is needed to evaluate their well-being in order to further assess and confirm the safety of the PGD technique, as proposed by several authors and the European Society of Human Reproduction and Embryology PGD Consortium (Vastdag, 2004; Klitzman, 2008; Harton *et al.*, 2011; Goldsammer and Jotkowitz, 2012).

The Universitair Ziekenhuis Brussel (UZ Brussel) is a centre with long experience in ART (since 1993) with one of the best prospectively followed cohorts of PGD children. We opted for two control groups to cover the complex reality of PGD. By comparing the PGD group with the ICSI group we were focusing closely on the effect of embryo biopsy. By comparing the PGD group with the SC group, we wanted to complete the picture by looking at the effect of embryo biopsy combined with ICSI treatment on fertile patients and their offspring. We therefore investigated whether preschool PGD children differ in their cognitive and motor development from children born after ICSI and SC in order to document the long-term psychological health outcomes after PGD and provide reassurance as to the safety of this technique.

Does PGD influence (i) the cognitive and (ii) the motor development of singletons at preschool age in comparison with children born after ICSI and SC?

Materials and Methods

Participants

Our sample consisted of three different groups: the PGD group, compared with an ICSI and a SC control group. PGD and ICSI children were recruited from an established cohort followed since birth. Their parents were invited by telephone. Given the extensive paediatric and neurological follow-up

programme for both ICSI and PGD, all children were investigated at the age of 2 months and 2 years. Neonatal and medical results for the cohort had been published previously (Desmyttere et al., 2012, 2009). ART parents and children were invited by telephone and letter. Recruitment of the SC children took place via different channels. Leaflets and letters were distributed in eight different kindergartens situated in the close vicinity of the hospital and up to 35 km away from it. Announcements were made via the hospital's intra and internet websites. Children were eligible if they were conceived naturally without any kind of medical intervention, singleton, Caucasian, Dutch speaking and born after 32 weeks of gestation. Multiples and very prematurely born children were excluded because of possible interference with their development (Miceli et al., 2000). Only Dutch speaking and Caucasians were included in order to overcome linguistic and socio-cultural barriers. The ICSI and SC children were case–controlled matched to the PGD children as closely as possible for gender, age, birth order, maternal educational level (low: partially completed school education or no qualifications at all, medium: school education completed, high: higher education qualification or a degree).

The technique used (aspiration of blastomeres) for the 8-cell embryos was the same in all PGD conceptions and identical to previous studies performed at our centre (Nekkebroeck et al., 2008; De Vos et al., 2009). The ICSI technique was equivalent in the ICSI group and the PGD group. The only technical difference between the ICSI and PGD groups was the embryo biopsy procedure. All embryos were transferred on Day 5.

Research protocol

All PGD and ICSI children were invited to the Centre of Medical Genetics for examination. First, a paediatrician carried out a medical examination. Afterwards, an experienced clinical psychologist (C.W.) introduced the test materials and the procedure to the children in the presence of their parents. Subsequently, the child was examined in the absence of his/her parents. At the end of the session, parents received verbal feedback regarding the assessment. The scored results of the intelligence test were sent home to the parents. In addition, all parents received standardized psychological questionnaires and inquiries concerning socio-demographic data, participant characteristics, details on pregnancy and neonatal data. For simplification, we will describe all these data below under the overarching term socio-demographic variables. All the instruments were provided in Dutch. Some of the parents filled out the forms during their child's assessment, while others received the questionnaires by e-mail. The analysis of these questionnaires capturing the children's social–emotional development will be reported elsewhere (Winter et al., in progress). The SC children were assessed in isolated classrooms in their kindergartens or if possible, the assessment took place at the Centre of Medical Genetics of UZ Brussel. In parallel, the SC children were assessed without their parents. The first author conducted all assessments, IQ and motor examinations.

For practical reasons, the assessor was not always blinded to the mode of conception. In the hospital setting, study nurses introduced the parents to the clinical psychologist without mentioning their mode of conception. Only in a few cases ($n = 5$) did parents spontaneously discuss their mode of conception/genetic disease. SC children, however, were mostly tested in kindergartens and so blinding was impossible.

Outcome measures

To measure the cognitive and psychomotor development of the 5- to 6 year olds, two standardized tests with acceptable to good psychometric qualities (i.e. validity and reliability) were consecutively administered. First, the Wechsler Preschool and Primary Scale of Intelligence III (WPPSI, Hendriksen and Hurks, 2010; Wechsler, 2009) was administered, followed by the Movement ABC (M ABC, Henderson and Sudgen, 1998). All instruments were provided in Dutch.

Cognitive development

The WPPSI-III-NL (Wechsler, 2009; Hendriksen and Hurks, 2010) is a widely used, individually administered, standardized measure of intelligence for children aged 2.6 years to 7.11 years. A verbal intelligence score (VIQ) and a performance intelligence score (PIQ) are combined, yielding a full-scale intelligence score (FSIQ). Raw scores were converted into scale scores by age and according to Flemish norms (Hendriksen and Hurks, 2010; Hurks 2010; Bos et al., 2011). Due to concerns related to time and attention span, we decided not to administer one optional subtest (i.e. the Symbol Search subtest) and so no measurement of processing speeds could be obtained. According to the Dutch Commission for testing matters (COTAN, 2010)—which certifies tests on their reliability and validity—the Dutch WPPSI-III-NL is qualified as having good criterion and construct validity and an acceptable degree of reliability.

Motor development

The M ABC (Henderson and Sudgen, 1998) is a standardized and widely used screening instrument for mild-to-moderate motor impairment (Watter, 2006; van Waelvelde et al., 2007). The test assesses the motor development of children aged 5–12 years via age-appropriate physical tasks comprising three subscales: manual dexterity, ball skills and static and dynamic balance. A total motor score can be calculated as the sum of the three subscales. Within these subscales, eight items can be rated on a six-point rating scale, where 5 represents the weakest and 0 represents the best performance. Profile scores provide more specific information on the child's movement skill performance in each individual category. A total impairment score expresses the child's test performance (Henderson and Sudgen, 1998; Cools et al., 2009).

Ethics

The Ethics Committee of the UZ Brussel approved this study. The parents were intensively informed beforehand. They all signed an informed consent form at the time of examination.

Statistical analyses

A priori power analysis showed that in order to detect medium-sized effects [i.e. a partial eta squared (η^2) value of 0.06] between our groups with 80% power, at least 152 children had to be included. This threshold was initially met.

Data analyses were performed using the Statistical Package for Social Sciences (SPSS) 20.0 for Windows. To detect possible confounders, two sets of exploratory analyses were performed. We followed the same statistical procedures for both outcome variables.

To detect between-group differences in data for PGD, ICSI and SC children, bivariate analyses were applied to all relevant socio-demographic variables (Table 1). Appropriate statistical tests were selected based on the measurement level of the data.

Within the groups, correlations between the WPPSI-III-NL or the M ABC and socio-demographic variables were computed using Pearson's r or Kendall's τ .

Before calculating the analysis of covariance (ANCOVA), we checked for violations of the underlying assumptions and controlled for outliers. Data were normally distributed for FSIQ and PIQ. Data for the VIQ were almost normally distributed with acceptable deviations. Moreover, Levene's test of homogeneity of variance showed that the variance across the groups was equal (data not shown).

Concerning the M ABC, the Levene's test of Equality of Error of Variances showed that the assumption of equal variances is not fulfilled ($F(5, 136) = 3.429, P = 0.006$). Furthermore, the distribution of the total scores for Movement showed a rather positive skew. Subtest analysis revealed that the distribution of ball skills and balance performances were also skewed

Table I Demographic parameters of PGD and ICSI children and those born after SC and their families.

	PGD (n = 47)	ICSI (n = 49)	SC (n = 48)	Test	P	Partial η^2 Phi
Male/female	17/30	21/28	20/28	$\chi^2(2) = 0.504$	0.777	0.059
Birth rank ^a						
Only child/eldest	9/18	9/20	8/20			
Middle position/youngest	2/18	2/18	7/18	$\chi^2(6) = 5.076$	0.534	0.188
Age at assessment in months ^b	64.47 ± 5.72	65.778 ± 5.92	67.15 ± 5.09	$F(2, 141) = 2.724$	0.069	0.037
Birthweight ^b (g)	3202 ± 514.58	3327 ± 515.78	3285 ± 445.14	$F(2, 141) = 0.802$	0.450	0.011
Gestational age in weeks	38.88 ± 1.85	39.18 ± 1.48	39.48 ± 1.74	$F(2, 141) = 1.491$	0.229	0.021
Gestation < 37 weeks	6/47	3/49	2/48	$\chi^2(2) = 2.731$	0.255	0.138
Admission NICU ^c						
1–7 days	2/47	4/49	3/48			
>8 days	3/47	0/49	3/48	$\chi^2(4) = 4.025$	0.403	0.167
Maternal age at birth ^b	30 ± 3.48	31.35 ± 3.91	30.73 ± 4.38	$F(2, 141) = 1.401$	0.250	0.019
Maternal educational level						
High/medium/low	26/20/1	29/18/1	36/11/1	$\chi^2(4) = 4.950$	0.292	0.185
Mean % of maternal employment ^b	72.45 ± 34.05	72.92 ± 35.74	79.79 ± 22.07	$F(2, 141) = 0.830$	0.438	0.012
Paternal age at test ^b	38.77 ± 4.48	40.18 ± 4.29	39.15 ± 5.34	$F(2, 141) = 1.0165$	0.315	0.016
Paternal educational level						
High/medium/low	22/22/2	20/28/1	27/19/2	$\chi^2(6) = 5.304$	0.505	0.192
Mean % of paternal employment ^b	92.98 ± 22.52	97.14 ± 14.72	96.60 ± 9.841	$F(2, 141) = 0.891$	0.412	0.13

Values are n unless otherwise stated. Education level: low: partially completed school education or no qualifications at all, medium: school education completed, high: higher education qualification or a degree.

^aBirth rank: four positions in the families are computed together.

^bValues are means ± SDs.

^cAdmission neonatal intensive care unit. 1–7 days and >8 days are computed together.

Table II Exclusion outliers of the WPPSI.

	Group	Gender	Gestation (weeks)	Maternal educational level	Full IQ	Verbal IQ	Performance IQ	Divergent subtest(s)
Child 1	ICSI	Girl	>37	Medium	130	129	111	Coding, 2.79
Child 2	SC	Girl	>37	High	115	108	120	Matrix reasoning, 2.82
Child 3	PGD	Boy	>37	Medium	105	106	104	Matrix reasoning, 3.17
Child 4	ICSI	Girl	>37	High	90	96	83	PIQ, 2.59
Child 5	SC	Girl	>37	Low	90	89	100	Vocabulary, 2.79
Child 6	PGD	Boy	>37	Medium	71	86	73	PIQ, 3.38; VIQ, 3.46

IQ, intelligence quotient; VIQ, verbal intelligence score; PIQ, performance intelligence score.

(data not shown). Despite some skewed distributions, ANCOVA has been found to be relatively robust in terms of deviations from normality (e.g. Glass *et al.*, 1972; Harwell *et al.*, 1992; Lix *et al.*, 1996).

To identify outliers, IQ and MABC scores were transformed into z-scores. Absolute values larger than the cut-off point of 2.5 (two and a half standard deviations) were excluded (Tables II and III). The choice of this cut-off point was based on power considerations. The initial choice of 2.0 as a cut-off point

resulted in the loss of too many cases. The same cut-off was used for all the outcome measures of the WPPSI-III-NL and the MABC. As Tables II and III demonstrate five (WPPSI) and six (Movement) cases, derived from all three conception groups, were therefore removed from the sample which decreased the power to detect a medium-sized effect.

An ANCOVA was carried out to test the significance of differences between the groups. Socio-demographic variables that showed significant

Table III Exclusion outliers of the M ABC.

	Group	Gender	Gestation (weeks)	Maternal educational level	Hand	Ball	Balance	Total	Divergent subtest(s)
Child 1	ICSI	Boy	>37	Medium	9	6	14.5	29.5	Ball: 2.67; balance: 5.07; total: 4.39
Child 2	PGD	Boy	>37	Medium	11	0	10	21	Balance: 3.26; total: 2.64
Child 3	ICSI	Boy	>37	High	6	6	4	16	Ball: 2.54
Child 4	ICSI	Girl	>37	High	9.5	5	6	20.5	Total: 2.54
Child 5	PGD	Boy	>37	Medium	11	0	9	20	Balance: 2.85

relationships in both the analyses mentioned above (i.e. they were related to group measures as well as our outcome measures) were used as covariates to answer our research question. These variables were age, mother's age at birth and educational level of both mother and father. Gender was included as a surplus fixed factor, as an extra control for differences between boys and girls. Dunnetts' T3 was used as an appropriate measurement for *post hoc* analysis due to the violation of the homogeneity of variance assumption. *P*-values <0.05 were deemed significant.

Inferential statistics, such as ANCOVA, require sufficient power to reach reliable and valid conclusions. Since we expect that the differences between our groups will be small, relatively large sample sizes are necessary in order to have enough statistical power. Since it is hard to find a sufficiently large sample in the population under study, we decided to provide additional statistical information. For each ANCOVA effect, a partial η^2 is reported. Partial η^2 is a standardized measure of effect size. It can vary between 0 and 1 and can be interpreted as the proportion of variance of the dependent variable that is explained by the independent variable. Cohen's (1988) guidelines were used for the interpretation of the effect: small effect size = 0.01, moderate effect size = 0.06 and large effect size = 0.14.

The proportion of variance in the outcome variable that is explained by the independent variable is reported for group membership (PGD versus ICSI versus control). This measure of effect size, however, is also subject to variability due to sampling and confidence intervals (CI) are therefore calculated for the effect size measures (Cumming, 2013). We therefore used the R Package MBESS to determine the 95% CIs of the effect sizes (Steiger, 2004). This CI defines a set of possible values for the effect size. Values outside the interval are relatively implausible. If an interval is short and relatively close to zero, one could conclude that the true value of the effect size is zero or very close to zero. One could then conclude that the effect is negligible. The upper bound of the CI can be regarded as the highest plausible value of the effect size. As such, we use the effect size itself as well as the upper bound of its CI to draw inferences about the clinical or practical importance of the effects found in this study.

Results

Participants and socio-demographic characteristics

Between January 2005 and May 2007, 65 PGD children met the inclusion criteria. This sample consisted of more girls ($n = 40$) than boys ($n = 25$), but this difference was not statistically significant. Between April 2011 and May 2013, 47 (72.3%) PGD children participated in our study and

were consequently assessed, while 10 (15.4%) were lost to follow-up and 8 (12.3%) refused. It is noteworthy that in this group of the 12.3% who refused, two children had been diagnosed with cerebral palsy. These children were in a revalidation programme and therefore could not join the study. A total of 68 ICSI children were contacted. Of these children, 4.41% ($n = 3$) had to be excluded due to insufficient, invalid test results (language barrier/conduct problems/illness during test situation), while 4.4% would not join, 11.8% were lost to follow-up and 7.3% could not participate for logistical reasons. Finally, 49 ICSI children took part in this study. As mentioned earlier, SC children were recruited via varying channels. Parents of ~300 children of eight different kindergartens were invited by letter to participate in this study. This resulted in 94 responders of whom 55 (58.51%) were eligible for participation and finally joined the study. Data of 7 children had to be excluded because of missing values; therefore, data of 48 (51.06%) SC children were finally analysed.

To confirm that the PGD group was still sufficiently representative, we compared the non-participant PGD group with our actual sample using three basic criteria (data not shown). The participants who did not respond, did not differ either in terms of gender $\chi^2(3, n = 170) = 0.46^*$, $P = 0.937$) or prematurity (born <37 weeks) $\chi^2(3, n = 170) = 5.73^*$, $P = 0.113$). Mothers with a lower educational level ($n = 4$) were represented significantly more in the non-participating group (Fisher's exact test: $P = 0.01$). Thus, families with lower level educational backgrounds were more often lost to follow-up ($n = 3$) or failure to attend child assessment ($n = 1$).

Differences between the conception groups regarding socio-demographic data were computed. As Table I shows children's characteristics such as age, birthweight, gestational age and admission to neonatal intensive care unit as well as parental characteristics did not show any significant difference when comparing the PGD group with the ICSI and the SC group. Table I also demonstrates that parents' educational level and employment percentage did not differ between the groups.

Cognitive development

Correlation coefficients between WPPSI-III-NL scores and socio-demographic background variables were computed to identify possible confounders. Mother's age at birth was significantly related to the FSIQ score of the WPPSI-III-NL ($r = 0.208$, $P = 0.012$), the PIQ

($r = 0.170$, $P = 0.042$) as well as the VIQ ($r = 0.191$, $P = 0.021$). The mother's level of education was related to the FSIQ ($\tau = -0.204$, $P = 0.002$), the VIQ ($\tau = -0.194$, $P = 0.004$) and the PIQ ($\tau = -0.222$, $P = 0.001$). The same correlations were detected for the father's educational level, FSIQ ($\tau = -0.157$, $P = 0.022$), the VIQ ($\tau = -0.168$, $P = 0.014$) and the PIQ ($\tau = -0.194$, $P = 0.005$). Other variables did not show any significant correlation.

ANCOVAs performed on the FSIQ, VIQ, PIQ data and subsequently for the subtests revealed no statistically significant differences between the groups (Table IV). Since group sizes were relatively small, we used exact CIs around the effect sizes to interpret the data and provide reassurance that the absence of a significant result was not solely due to insufficient power of this study. This revealed with 95% reliability that the effect of the conception group on the measurements of the FSIQ would not account for >4.3% of the variance if the study were to be repeated.

The VIQ and its subscales yielded no between-group differences (Table IV). Concerning the VIQ, the mother's level of education was the only significant covariate, but also lacked substantial relevance in terms of effect size ($F(1, 134) = 5.289$, $P = 0.023$, $\eta^2 = 0.038$) ($P = 0.023$, $\eta^2 = 0.038$). Performances on the PIQ were similar among the conception groups (Table IV). As for the PIQ, ANCOVA revealed that the maternal educational level was significantly influential ($F(1, 135) = 5.163$, $P = 0.025$, $\eta^2 = 0.037$), but again the small effect size indicated clinical irrelevance.

All the subtests within the performance scale also revealed no group differences. Controlling for confounders on subtest levels did suggest

some influences, but partial η^2 revealed that these did not imply any clinical importance (data not shown).

We conclude that the cognitive performance of PGD children did not differ from that of ICSI or SC children.

Motor development

Correlation coefficients between the M ABC and its subscales and the socio-demographic data were computed. The age of the children correlated with manual dexterity skills ($r = -0.173$, $P = 0.035$), balance ($r = -0.162$, $P = 0.048$) and Movement Total ($r = -0.197$, $P = 0.016$). These correlations show that as children grow older their performance improves. The average percentage of time fathers worked correlated weakly with manual dexterity skills ($r = 0.174$, $P = 0.038$). The mother's educational level was weakly related to performance for balance skills ($r = 0.150$, $P = 0.035$) and to Movement Total performance ($r = 0.144$, $P = 0.033$).

Finally, ANCOVAs were applied to Movement ABC scores (Table V). Again covariates were chosen based on their significant correlations with the Movement Total score and the subtest scores.

As Table V demonstrates the ANCOVA for the Movement Total, test scores revealed a significant difference between the groups. *Post hoc* analysis with Dunnett's T3 test ($t = 0.927$; $P = 0.009$) indicated that SCs ($M = 7.03$, $SD = 3.94$) performed better than ICSIs ($M = 9.85$, $SD = 5.21$). ANCOVAs demonstrated significant gender differences on global Movement performance ($F(1, 133) = 12.072$,

Table IV Data analysis of the WPSI of PGD children versus ICSI children and SC children.

	PGD (n = 47)	ICSI (n = 49)	SC (n = 48)	ANOVA	P	Partial η^2	95% CIs
Full IQ	117.21 ± 13.25	115.55 ± 14.41	118.85 ± 12.70	$F(2, 134) = 0.436$	0.647	0.006	[0, 0.43]
VIQ	116.77 ± 11.58	113.35 ± 12.95	115.63 ± 10.96	$F(2, 134) = 1.370$	0.258	0.020	[0, 0.74]
Information	12.57 ± 2.65	11.78 ± 2.62	12.33 ± 2.38	$F(2, 134) = 1.542$	0.218	0.023	
Vocabulary	13.19 ± 2.12	12.82 ± 2.26	12.90 ± 2.02	$F(2, 134) = 0.988$	0.375	0.015	
Word reasoning	13.51 ± 2.38	12.61 ± 2.91	13.10 ± 2.53	$F(2, 134) = 1.671$	0.192	0.024	
PIQ	115.62 ± 10.58	114.34 ± 14.13	114.34 ± 14.13	$F(2, 134) = 0.462$	0.631	0.007	[0, 0.004]
Block Design	12.38 ± 2.92	11.82 ± 2.95	13.15 ± 3.06	$F(2, 134) = 1.376$	0.256	0.020	
Matrix reasoning	12.55 ± 2.46	12.35 ± 2.04	12.46 ± 2.48	$F(2, 134) = 0.023$	0.997	0.000	
Picture concept	12.23 ± 2.07	12.53 ± 3.26	12.56 ± 2.63	$F(2, 134) = 0.603$	0.549	0.009	
Coding	11.02 ± 2.22	10.86 ± 2.73	11.38 ± 2.42	$F(2, 134) = 0.929$	0.397	0.014	

Means and SDs are reported Full IQ = VIQ plus PIQ. CI, confidence interval, ANOVA, analysis of variance.

Table V Data analysis of the M ABC performances of children born after PGD versus children born after ICSI or SC.

	PGD (n = 47)	ICSI (n = 49)	SC (n = 48)	ANOVA	P	Partial η^2	95% CI
Total Movement	7.51 ± 4.99	9.85 ± 5.21	7.03 ± 3.94	$F(2, 133) = 3.49$	0.033	0.050	[0, 0.124]
Manual dexterity score	4.57 ± 3.04	5.19 ± 2.47	4.76 ± 2.30	$F(2, 133) = 0.14$	0.868	0.002	[0, 0.024]
Ball skill score	1.17 ± 1.32	1.88 ± 2.03	1.14 ± 1.61	$F(2, 133) = 2.97$	0.055	0.043	[0, 0.113]
Balance skill score	1.76 ± 2.49	2.87 ± 2.88	1.09 ± 1.63	$F(2, 133) = 5.72$	0.004	0.079	[0.025, 0.163]

Values are means and SDs.

$P = 0.001$, $\eta^2 = 0.083$), with girls ($M = 6.91$; $SD = 3.99$) doing better than boys ($M = 10.03$; $SD = 5.52$) in terms of a medium-sized effect.

There was a trend for ball skills to differ significantly by group (Table V). *Post hoc* analysis using Dunnett's T3 did not confirm a significant difference. As we can see in Table V balance skills scores differed significantly between the three conception groups. *Post hoc* Dunnett's T3 test ($t = 0.473$; $P = 0.002$) demonstrated that SC children did better on balance tasks than ICSI children (Table V). Boys and girls differed in terms of balance performance ($F(1, 133) = 9.28$, $P = 0.003$, $\eta^2 = 0.065$): yielding better scores for females ($M = 1.37$; $SD = 1.82$) than for males ($M = 2.76$; $SD = 3.12$) with an effect of medium power.

No significant group differences could be traced in the manual dexterity (Table V). Gender also had an effect on the results for manual dexterity ($F(1, 133) = 19.63$, $P = 0.000$, $\eta^2 = 0.129$): girls were significantly better ($M = 4.13$, $SD = 0.26$) than boys ($M = 5.98$; $SD = 0.32$); the size of this effect is rather large at 12.9%.

These results preclude the conclusion that PGD children's motor performance is equivalent to that of SC and ICSI children.

Discussion

To our knowledge this is the first study to investigate whether preschool PGD children differ significantly in their cognitive and motor development in comparison with children born after ICSI or SC. The most important finding is that the mode of conception had no significant influence on cognitive development. In terms of motor development, no significant differences between the performance of PGD children and both control groups could be identified in our sample. However, ICSI children performed less well than SC children especially on balance tasks. Girls did better on some subtests than boys.

Since this was the first study to compare an exclusive PGD group of pre-schoolers to matched controls, we cannot refer to similar studies. Nonetheless, our results are in line with findings described in a mixed PGD/PGS group (Banerjee et al., 2008; Nekkebroeck et al., 2008). Thomaidis et al. (2012) additionally pointed out that 77.4% of the PGD children investigated showed normal general development. The latter study, however, lacked a matched control group.

Since the WPSSI-III-NL is a standardized and widely used instrument, it allows us to compare our cognitive test results to norms for the Dutch-speaking (i.e. Flemish) population of Belgium. The majority of our participants' scores in the three groups were above average. It is well established that ART couples are more frequently more highly educated than the general population (Place and Englert, 2003). On the other hand, a large proportion of the parents with a participating spontaneously conceived child wanted their impression of their children's intelligence to be confirmed with an IQ test. This effect has been described before by others (Place and Englert, 2003; Knoester et al., 2008).

The motor development of the PGD children in our study suggests that PGD children do not differ significantly from the two control groups. This was the case for total scores and also for their performance on the subtests.

Looking at this in greater detail, SC children had better results than ICSI children for the total Movement score as well as for ball and balance skills.

Our results are essentially in line with the findings of Nekkebroeck et al. (2008) and Schendelaar et al. (2013), while Banerjee et al. (2008) and Thomaidis et al. (2012) found rather delayed motor scores, while only the latter focused exclusively on PGD children. As already

mentioned, these studies used different tests and they included twins and triplets (who are more prone to neonatal complications and therefore delayed development). However, Banerjee et al. (2008) included control twins who were matched as closely as possible to the PGD/PGS group. Here the incidence of neonatal adversity differed significantly between PGD/PGS children and controls, but the authors did not mention whether multiplicity accounted for cognitive differences.

Though the developmental outcomes of ICSI children are outside the scope of the present paper, it is noteworthy that the motor performance of ICSI children is only lower in comparison with SC children and not in comparison with PGD children. As Zhu et al. (2009, 2010) have pointed out, the low fertility of ICSI parents may play a certain role here. Our results are in line with the slightly poorer motor skills of ICSI children found in several studies (Bouwen et al., 1998; Sutcliffe et al., 1999; Belva et al., 2007; Knoester et al., 2007; Zhu et al., 2009, 2010).

Although the effect sizes found in our sample were rather small, the CIs indicate that the effect in the population could range from zero to a rather large effect (with the upper limit for the effect size of balance capacities being 16.3%). This means that our results should therefore be interpreted with some caution.

We also noticed skewed distributions and variability in the performance seen on the M ABC. This could be due to several factors. First, age-specific changes such as performance in the speed of repetitive tasks (as in the subtest manual dexterity) (Largo et al., 2003) as researchers have observed a greater inter-individual variation within the same age group (Largo et al., 2003). Secondly, the M ABC has been criticized for loss of specificity and high variability in general (Watter, 2006; Cools et al., 2009). Thirdly, Largo et al. (2003) described that gender differences may occur and vary in size and direction depending on the tasks. This is consistent with our results: gender differences played a role in the children's performance on all levels. In ball tasks girls and boys performed adequately, while manual dexterity and balance tasks favoured girls. Our findings mirror the results of Leunens et al. (2006) who showed that 8-year-old ICSI girls had significantly better manual skills than boys, whereas 10-year-old girls had better balance skills than their male counterparts (Leunens et al., 2008). We cannot, however, ignore the fact that factors such as fatigue, hunger and motivation may have played a role.

Strengths and limitations

Our study is the first case–controlled, matched investigation comparing exclusively PGD singletons with two control groups. This is undoubtedly a strong point given the fact that PGD and ICSI share the same technique but these two ART groups need also an appropriate group of spontaneous controls. Other authors also include PGS children or have no controls at all (Banerjee et al., 2008; Nekkebroeck et al., 2008; Thomaidis et al., 2012). In the absence of adequate controls, the effect of subfertility remains unclear since PGS parents often have genetic problems combined with subfertility. As Carson et al. (2010) suggested, it is of vital importance to differentiate what the aim of measurement is. We therefore opted to match the PGD group with two control groups covering the complex reality of PGD. By comparing the PGD with the ICSI group, we shed light on the possible effect of the embryo biopsy since this is the main difference between ICSI and PGD. By comparing the PGD group with the SC children, we focused on the combined effect of the fertility treatment and the genetic background. The advantage of our

method is its efficiency in controlling for confounding factors. Its limitation, however, leads to the exclusion of valuable data such as premature children. It is also noteworthy that two children with cerebral palsy from the original sample were finally excluded from our study. Other authors (Middelburg *et al.*, 2011; Schendelaar *et al.*, 2013) also described cerebral palsy after embryo biopsy; however, in a PGS population.

Another limitation is that matched controlled studies are time and labour intensive but may lack sufficient statistical power (Sutcliffe *et al.*, 1995; Carson *et al.*, 2010). Our power analysis suggested that to detect at least a medium effect (i.e. a partial η^2 value of 0.06) with 80% power using ANCOVA, at least 52 children had to be included per group. By removing the outliers and excluding cases because of missing values, we fell short of this threshold and have therefore reduced the power of our analyses.

The small sample size may also exclude, for example, children diagnosed with attention deficit hyperactivity disorder or autism spectrum disorders. Those can easily be missed in samples of this size. We must recognize, however, that diagnosis of these severe developmental disorders is very complex and demands experienced professionals (especially child psychiatrists). Such a goal was beyond the scope of this project. National register-based studies do indeed cover a much broader population and can easily detect small increased risks, such as tic disorders (Bay *et al.*, 2013) or autism and mental retardation (Sandin *et al.*, 2013) among ICSI children. Belgium, however, is one of the few countries in Europe that practice PGD with longstanding experience but has no national health registers. Furthermore, it is worth mentioning that the PGD group is still a small group in comparison with other ART techniques. Only a maximum of 65 PGD children could be included in our study.

It is a shortcoming that a definite attrition/participation rate could not be calculated. Our participation rate for the PGD group is moderate at 71.9%, acknowledging that most non-participants were those that could not be reached. Obviously, this precludes the calculation of a precise participation rate. Omitting those lost to follow-up, our participation rate increases from 71.9 to 87.5%. Our sample is representative with reference to prematurity (born at 32–36 weeks) and gender. Families with lower educational backgrounds, however, drop out more quickly from long-term follow-up programmes (Belva *et al.*, 2012). Also, bias within the control group is possible since participation was completely voluntary and the children recruited may have participated because of an ascertainment bias (parents who think that their child is very intelligent or on the other hand who think that their child might have a special problem, are more willing to join a study (Middelburg *et al.*, 2008). Since announcements for the control group were widely disseminated via multiple channels, many parents should have had received our invitation: 94 of these responded and 51.06% of the eligible SC children joined.

In addition, investigations of PGD and ICSI children were carried out in the same setting by a single administrator who was mostly blinded to the mode of conception for the ART groups.

The ART group assessments took place in the hospital, however, whereas the majority of the SC group was assessed in kindergartens, which might be a source of bias.

The proximity of parents in the ART groups may have given the PGD/ICSI children an advantage, e.g. feeling more secure. The hospital environment itself may, however, have been a disadvantage, causing extra stress. This was counterbalanced in turn by the fact that SC children

were embedded in the daily routine of kindergarten activities; the assessment might have been seen as a special game or they may have felt interrupted in their play.

Given that multiples and non-Caucasians were excluded from the study and only a few children were born prematurely, our findings may not be generalized to all children born after PGD. We would therefore suggest that future studies should pay attention to premature PGD children, multiples and diverse cultural groups. As in previous follow-up studies, research should include the social–emotional development of children born after PGD and context-related variables, such as family characteristics and evaluations of their environment, to broaden the focus on development. This is especially important within a group of children whose parents may have a higher potential for traumatization due to genetic conditions in their families of origin.

To circumvent issues related to statistical power, mainly because of the limited sample size, we opted to calculate CIs around the measure of effect size. In most analyses, the upper limit of this CI indicated that, if the study were replicated, the effect sizes would still be small (WPPSI-III NL).

For motor measurements; however (M ABC), the CIs around effect sizes are rather large. Although we chose the M ABC for practical reasons (easy to administer within a reasonable time span and with minimal burden to pre-schoolers), as a standardized and widely used screening tool for motor impairment we acknowledge that this test has its limitations in terms of sensitivity and specificity (Cools *et al.*, 2009). Replication of motor measurements at a later age using a more reliable instrument seems to be warranted.

In conclusion, PGD singletons have similar cognitive capabilities compared with SC and ICSI singletons. With respect to motor performance, our data suggest that PGD singletons perform as well as controls. Long-term follow-up at an older age, with less intra-individual and age-related differences in motor development, will be mandatory in order to confirm our results. These results call for confirmatory investigations and long-term follow-up in multicentre cohort studies.

Acknowledgements

We want to thank children and parents for their time and voluntary cooperation, Andrea Buysse and Leen Ausloos for practical help and support in conducting the study. Special thanks to Florence Belva for commenting on the manuscript draft.

Authors' roles

C.W. matched the children, invited the families, tested the children, performed the statistical analyses and wrote the article. S.D. and F.D.S. carried out medical examinations of the children. F.V.A. made substantial contributions with regard to data acquisition, analysis and interpretation of the data. J.N., S.D. and M.B. were responsible for the study design and supervision. All authors were involved in drafting/revision of this paper and read the manuscript before submission.

Funding

This study was funded by the Onderzoeksraad (OZR) of the Vrije Universiteit Brussel and the FWO (Fonds Wetenschappelijk Onderzoek) and Wetenschappelijk Fonds Willy Gepts.

Conflict of interest

U.Z.B. and the Centre for Medical Genetics have received many educational grants for organizing the data collection, from IBSA, Ferring, Organon, Shering-Plough, Merck and Merck Belgium. M.B. has received consultancy and speaker's fees from Organon, Serono Symposia and Merck.

References

- Banerjee I, Shevlin M, Taranissi M, Thornhill A, Abdalla H, Ozturk O, Barnes J, Sutcliffe A. Health of children conceived after preimplantation genetic diagnosis: a preliminary outcome study. *Reprod Biomed Online* 2008; **16**:376–381.
- Basille C, Frydman R, Abdelwahab El A, Hesters L, Fanchin R, Tachdjian G, Steffann J, LeLorc'h M, Achour-Frydman N. Preimplantation genetic diagnosis: state of the art. *Eur J Obstet Gynecol Reprod Biol* 2009; **145**:9–13.
- Bay B, Mortensen EL, Hvidtjørn D, Kesmodel U. Fertility treatment and risk of childhood and adolescent mental disorders: register based cohort study. Fertility treatment and risk of childhood and adolescent mental disorders: register based cohort study. *BMJ* 2013; **347**:1–11. doi:10.1136/bmj.f3978.
- Belva F, Henriët S, Liebaers I, Van Steirteghem A, Celestin-Westreich S, Bonduelle M. Medical outcome of 8-year-old ICSI children (born > = 32 weeks' gestation) and a spontaneously conceived comparison group. *Hum Reprod* 2007; **22**:506–516.
- Belva F, Roelants M, Painter R, Bonduelle M, Devroey P, De Schepper J. Pubertal development in ICSI children. *Hum Reprod* 2012; **27**:1156–1161.
- Bos A, Magez W, Schittekatte M. De WPSSI-R en de WPSSI II in de focus. *Signal* 2011; **2**:38–42.
- Bouwen 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.
- Carson C, Kurinczuk JJ, Sacker A, Kelly Y, Klemetti R, Redshaw M, Quigley MA. Cognitive development following ART: effect of choice of comparison group, confounding and mediating factors. *Hum Reprod* 2010; **1**:244–252.
- Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Erlbaum, 1988.
- Cools W, De Martelaer K, Samaey C, Andries C. Movement skill assessment of typically developing preschool children. A review of seven movement skill assessment tools. *J Sport Sci Med* 2009; **8**:154–168.
- COTAN. *Documentatie aan Tests en Testresearch in Nederland*. Amsterdam: Boom, test uitgevers, 2010.
- Cumming G. The new statistics: why and how. *Psychol Sci* 2013. doi:10.1177/0956797613504966.
- Desmyttere S, de Schepper J, Nekkebroeck J, De Vos A, De Rycke M, Staessen C, Liebaers I, Bonduelle M. Two year auxological and medical outcome of singletons born after embryo biopsy applied in preimplantation genetic diagnosis or preimplantation screening. *Hum Reprod* 2009; **24**:470–476.
- Desmyttere S, De Rycke M, Staessen C, Liebaers I, De Schrijver F, Verpoest W, Haentjens P, Bonduelle M. Neonatal follow-up of 995 consecutively born children after embryo biopsy for PGD. *Hum Reprod* 2012; **27**:288–293.
- De Vos A, Staessen C, De Rycke M, Verpoest W, Haentjens P, Devroey P, Liebaers I, Van de Velde H. Impact of cleavage-stage embryo biopsy in view of PGD on human blastocyst implantation: a prospective cohort of single embryo transfers. *Hum Reprod* 2009; **24**:2988–2996.
- Fisher-Jeffes LJ, Banerjee I, Sutcliffe AG. Parents' concerns regarding their ART children. *Reproduction* 2006; **131**:389–394.
- Glass GV, Peckham PD, Sanders JR. Consequences of failure to meet assumptions underlying fixed effects analyses of variance and covariance. *Rev Educ Res* 1972; **42**:237–288.
- Goldsammler M, Jotkowitz A. The Ethics of PGD: what about the physician? *Am J Bioethics* 2012; **12**:28–29.
- Handyside AH. Preimplantation genetic diagnosis after 20 years. *BioMed Online*, 2010; **21**:280–282.
- Harper JC, Wilton L, Traeger-Syndinos J, Goossens V, Moutou C, SenGupta SB, Pehlivan Budak T, Renwick P, De Rycke M, Geraerds JPM et al. The ESHRE PGD Consortium: 10 years of data collection. *Hum Reprod Update* 2012; **18**:234–247.
- Harton G, Braude P, Lashwood A, Schmutzler A, Traeger-Syndinos J, Wilton L, Harper JC; European Society for Human Reproduction (ESHRE) PGD Consortium. ESHRE PGD consortium best practice guidelines for organization of a PGD centre for PGD/preimplantation genetic screening. *Hum Reprod* 2011; **26**:14–24.
- Harwell MR, Rubinstein EN, Hayes WS, Olds CC. Summarizing Monte Carlo results in methodological research: the one- and two-factor fixed effects ANOVA cases. *J Educ Stat* 1992; **17**:315–339.
- Henderson SE, Sudgen DA. *Movement Assessment Battery for Children. Handleiding*. Lisse, The Netherlands: Swets and Zeitlinger B.V., 1998.
- Hendriksen J, Hurks P. *WPSSI III NL, Aframe en Scoringshandleiding*. Amsterdam: Pearson Assessment and Information BV, 2010.
- Hurks P. De nieuwe Wechsler kleuterintelligentietest voor 2 1/2 tot 8-jarigen. *Tijdschrift Voor Neuropsychologie* 2010; **2**:40–50.
- Izat Y, Goldbeck L. Die Entwicklung von Kindern aus assistierten Befruchtungen. Eine Übersicht der Studienlage. *Prax Kinderpsychol Kinderpsychiat* 2008; **57**:264–281.
- Klitzman R. Anticipating issues related to increasing preimplantation genetic diagnosis use: a research agenda. *BioMed Online* 2008; **17**:33–42.
- Knoester M, Helmerhorst FM, van der Westerlaken LA, Walther FJ, Veen S, Leiden Artificial Reproductive Techniques Follow-up Project (L-art-FUP). Matched follow-up study of 5 8-year-old ICSI singletons: child behaviour, parenting stress and child (health-related) quality of life. *Hum Reprod* 2007; **22**:3098–3107.
- Knoester M, Helmerhorst FM, Vandenbroucke JP, van der Westerlaken LAJ, Walther FJ, 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.
- Largo RH, Fischer JE, Rousson V. Neuromotor development from kindergarden age to adolescence developmental course and variability. *Swiss Med Weekly* 2003; **133**:193–199.
- 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 children born after ICSI compared to spontaneously conceived children. *Hum Reprod* 2008; **23**:105–111.
- Liebaers I, Desmyttere S, Verpoest W, De Rycke M, Staessen C, Sermon K, Devroey P, Haentjens P, Bonduelle M. Report on a consecutive series of 581 children born after blastomere biopsy for preimplantation genetic diagnosis. *Hum Reprod* 2010; **25**:275–282.
- Lix LM, Keselman JC, Keselman HJ. Consequences of assumption violations revisited: a quantitative review of alternatives to the one-way analysis of variance F test. *Rev Educ Res* 1996; **66**:579–619.
- Ludwig A, Katalinic A, Thyen U, Sutcliffe AG, Dietrich K, Ludwig M. Neuromotor development and mental health at 5.5 years of age of singletons born after intracytoplasmic sperm injection ICSI: results of a

- prospective controlled single blinded study in Germany. *Fertil Steril* 2009; **91**:125–132.
- Miceli PJ, Goeke-Morey MC, Whitman TL, Kolberg KS, Miller-Loncar C, White RD. Brief report: birth status, medical complications and social environment: individual differences in development of preterm, low birth, weight infants. *J Pediatr Psychol* 2000; **25**:353–358.
- Middelburg KJ, Heineman MJ, Bos AF, Hadders-Algra M. Neuromotor, cognitive, language and behavioural outcome in children born following IVF or ICSI—a systematic review. *Hum Reprod Update* 2008; **14**:219–231.
- Middelburg KJ, van der Heide M, Houtzager B, Jongbloed-Pereboom M, Fidler V, Bos AF, Kok J, Hadders-Algra M. PGS Follow-up Study Group. Mental, psychomotor, neurologic, and behavioral outcomes of 2-year-old children born after preimplantation genetic screening: follow-up of a randomized controlled trial. *Fertil Steril* 2011; **96**:165–169.
- Nekkebroeck J, Bonduelle M, Desmyttere S, Van den Broeck W, Ponjaert-Kristoffersen I. Mental and psychomotor development of 2-year-old children born after preimplantation genetic diagnosis/screening. *Hum Reprod* 2008; **23**:1560–1566.
- 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-B, Tarlatzis BC, Peters C, Hageberg BS, 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.
- R Package MBESS. <http://www3.nd.edu/~kkelley/site/MBESS.html>.
- Sandin S, Nygren K-G, Iliadou A, Hultman CM, Reichenberg A. Autism and mental retardation among offspring born after in vitro fertilization. *JAMA* 2013; **310**:75–84. doi:10.1001/jama.2013.7222.
- Schendelaar P, Middelburg KJ, Bos AF, Heineman MJ, Kok JH, La Bastide-Van Gemert S, Seggers J, van den Heuvel ER, Hadders-Algra M. The effect of preimplantation genetic screening on neurological, cognitive and behavioural development in 4-year-old children: follow-up of RCT. *Hum Reprod* 2013; **28**:1508–1518.
- Sermon K, Van Steirteghem A, Liebaers I. Preimplantation genetic diagnosis. *Lancet* 2004; **363**:1633–1641.
- Staessen C, Platteau P, Van Assche E, Michiels A, Tournaye H, Camus M, Devroey P, Liebaers I, Van Steirteghem A. *Hum Reprod* 2004; **19**:2849–2858.
- Steiger JH. Beyond the F test: effect size confidence intervals and tests of close fit in the analysis of variance and contrast analysis. *Psychol Methods* 2004; **2**:164–182.
- Sutcliffe AG, D'Souza SW, Cadman J, Richards B, McKinlay IA, Lieberman B. Minor congenital anomalies, major congenital malformations and development in children conceived from cryopreserved embryos. *Hum Reprod* 1995; **10**:3332–3337.
- Sutcliffe AG, Ludwig M. Outcome of assisted reproduction. *Lancet* 2007; **37**:351–359.
- Sutcliffe AG, Taylor B, Li J, Thornton S, Grudzinkas JG, Lieberman BA. Children born after intracytoplasmic sperm injection: population control study. *BMJ* 1999; **318**:704–705.
- Sutcliffe AG, Saunders K, McLachlan R, Taylor B, Edwards Ph, Grudzinkas JG. Outcome in the second year of life after in-vitro fertilization by intracytoplasmic sperm injection: a UK case control study. *Lancet* 2001; **357**:2080–2084.
- Thomaidis L, Kitsiou-Tzli S, Critselis E, Drandakis H, Touliatou V, Mantoudis S, Leze E, Destouni A, Traeger-Syndinos J, Kafetzis D et al. Psychomotor development of children born after preimplantation genetic diagnosis and parental stress evaluation. *World J Pediatr* 2012; **8**:309–316.
- Van Waelvelde H, Peersman W, Lenoir M, Smits Engelsman BCM. The reliability of the movement assessment battery for children for preschool children with mild to moderate motor impairment. *Clin Rehabil* 2007; **21**:465–470.
- Vastdag B. Group calls for stricter rules for assisted reproduction, ban of 'extreme' technologies. *JAMA* 2004; **291**:2306–2308.
- Verlinsky Y, Cohen J, Munne S, Gianaroli L, Simpson JL, Ferrarett AP, Kuliev A. Over a decade of experience with preimplantation genetic diagnosis: a multicenter report. *Fertil Steril* 2004; **82**:292–294.
- Verpoest W, Haentjes P, De Rycke M, Staessen C, Sermon K, Bonduelle M, Devroey P, Liebaers I. Cumulative reproductive outcome after preimplantation genetic diagnosis: a report on 1498 couples. *Hum Reprod* 2009; **24**:2951–2959.
- Watter P. Movement assessment battery of children. Commentary. *Aust J Physiother* 2006; **52**:68.
- Wechsler D. *Wechsler Preschool and Primary Scale of Intelligence III*. Amsterdam: NCS Pearson, INC., 2009.
- Zhu JL, Basso O, Obel C, Hvidtjærn D, Olsen J. Infertility, infertility treatment and psychomotor development: the Danish National Birth Cohort. *Paediatr Perinat Epidemiol* 2009; **23**:98–106.
- Zhu JL, Obel C, Basso O, Olsen J. Parental infertility and developmental coordination disorder in children. *Hum Reprod* 2010; **24**:908–913.