Neurodevelopmental outcome of children born following assisted reproductive technology

Middelburg, Karin Janette

Part III - Follow-up of children born after IVF with Preimplantation Genetic Screening

8

Mental, psychomotor, neurological, and behavioural outcome of 2-year-old children born following preimplantation genetic screening: follow-up of a randomised controlled trial

K.J. Middelburg
M. van der Heide
B. Houtzager
M. Jongbloed-Pereboom
V. Fidler
A.F. Bos
J.H. Kok
M. Hadders-Algra
for the PGS follow-up study group

Fertility and Sterility 2011; 96: 165-169.

ABSTRACT

Objective: To evaluate the effect of preimplantation genetic screening (PGS) on neurodevelopmental outcome in children.
Design: Prospective, assessor-blinded, follow-up study of children born to women randomly assigned to IVF/ICSI with or without PGS.
Setting: University Medical Center, Groningen, and Academic Medical Center, Amsterdam, The Netherlands.
Patients: Fifty-four PGS-children and 77 controls.
Intervention: PGS.
Main outcome measure(s): Mental, psychomotor, neurological and behavioural outcome in two-year-old children measured with the Bayley Scales of Infant Development, the Hempel neurological examination, and the Child Behaviour Check List.
Results: Mental, psychomotor, and behavioural outcome at 2 years in children born following IVF with and without PGS was overall similar. PGS-children showed lower neurological optimality scores than control children. Scores on all tests were within the normal range.
Conclusions: PGS-conception does not seem to be associated with impaired mental, psychomotor, and behavioural outcome at age 2. The lower neurological optimality scores found in PGS-children may signal less favourable long-term neurological outcome in PGS-children. Our findings stress the need of evaluation of safety of new assisted reproductive techniques before large-scale implementation.

INTRODUCTION

In preimplantation genetic screening (PGS), embryos are screened for chromosome aneuploidies with the intent to improve ongoing pregnancy rates after in vitro fertilisation (IVF). In this procedure, the zona pellucida of the embryo is opened with laser or by chemical means so that one or two blastomeres can be removed by biopsy. These blastomeres are then screened for aneuploidy and in this way embryos with an abnormal chromosomal constitution are identified and discarded. Theoretically, selection of the most viable embryos with PGS leads to higher ongoing pregnancy rates (Wilton, 2002). Results of recent randomised controlled trials, however, were unable to confirm this theory (Staessen *et al.*, 2004;Jansen *et al.*, 2008;Schoolcraft *et al.*, 2008;Staessen *et al.*, 2008;Meyer *et al.*, 2009;Debrock *et al.*, 2010). In fact, a reduction in ongoing pregnancies after IVF with PGS has been reported (Mastenbroek *et al.*, 2007;Hardarson *et al.*, 2008).

The invasiveness of PGS necessitates careful follow-up of children born after this procedure. So far, few controlled studies have investigated developmental outcome of PGS-children. Previously, our group reported on developmental outcome during infancy of children born at one centre of a two-centre randomised controlled trial on the effects of PGS. The small study reported no statistically significant differences in neurological outcome between children born after IVF with PGS and those born after IVF without PGS up until the age of eighteen months (Middelburg *et al.*, 2010). However, dysfunction in fine motor abilities and mild dysfunctions in posture and muscle tone regulation tended to be present more frequently following PGS. The small sample size of the study precluded the conclusion that PGS is not associated with a less favourable neurological outcome. So far, only two other groups reported on children born after PGS. They reported similar developmental and behavioural scores in children born following PGS or natural conception (Banerjee *et al.*, 2008) and found no adverse effect of PGS on growth, congenital malformations, neonatal intensive-care admissions, behaviour, and mental and psychomotor development (Nekkebroeck *et al.*, 2008a;Nekkebroeck *et al.*, 2008b;Desmyttere *et al.*, 2008).

The present study evaluates neurodevelopmental outcome of 2 year old PGS-children in comparison with children conceived after conventional IVF. To this aim, we conducted a prospective assessor-blinded follow-up study of children born to women participating in a multicentre randomised controlled trial (RCT) on IVF with or without PGS (Mastenbroek *et al.*, 2007).

METHODOLOGY

Eligibility and recruitment of participants

Eligible for the present follow-up study were all children born to women participating in a double blind RCT on the efficiency of PGS to improve ongoing pregnancy rates after IVF (Mastenbroek *et al.*, 2007). Inclusion criteria for women participating in the trial were: age 35 to 41 years, no previously failed IVF-cycles, and no objections against a possible double embryo transfer. Randomisation of women was performed centrally with minimization for age (35-37 or 38-41 years) and reproductive technique (IVF or ICSI), with stratification according to study centre (University Medical Center Groningen (UMCG) and Academic Medical Center (AMC), Amsterdam) prior to the start of the IVF-procedures. Further details on study design and information concerning IVF-treatment procedures have been reported previously (Mastenbroek *et al.*, 2007;ISRCTN76355836).

Before inclusion and randomisation, couples were informed that a follow-up program was part of the PGS-trial and that children born to couples who were included in the PGS-trial would be invited for neurodevelopmental follow-up.

The protocol of the follow-up study was approved by the Dutch Central Committee of Research Involving Human Subjects and the Medical Ethics Committees of the local hospitals. For the children participating in the follow-up study, parental written informed consent was provided.

Outcome measures

The Dutch standardization of the Bayley Scales of Infant Development - second edition (BSID-II-NL; (Van der Meulen *et al.*, 2004)) was used to evaluate mental and psychomotor development. The mental developmental index (MDI) is determined by performance on items measuring visual and auditory information processing, memory, language development, eye-hand coordination, imitation, and problem-solving abilities. The psychomotor developmental index (PDI) consists of items measuring fine and gross-motor skills. Dutch standardization norms (mean ± standard deviation: 100 ±15) were applied (Van der Meulen *et al.*, 2004). The BSID-II-NL mental and psychomotor scales were administered by trained psychologists.

Neurological function was assessed with the standardised neurological examination according to Hempel (Hempel, 1993), an assessment which pays special attention to the presence of minor neurological dysfunction (MND). In this assessment, fine-motor function, gross-motor function, posture and muscle tone, reflexes, and visuomotor function are tested during standardised play-sessions. Multiple signs in one of these domains result in a so-called dysfunctional domain (Hadders-Algra, 2003). In case of the absence of a dysfunctional domain or only the

presence of dysfunctional reflexes, children are classified as neurologically normal. The isolated presence of fine motor, gross motor, or visuomotor dysfunction or mild dysregulation of posture and muscle tone is defined as simple MND. It reflects a normal, but non-optimal form of brain function (Hadders-Algra, 2003). When two or more domains of dysfunction are present, a child is classified as having complex MND. This type of MND is clinically relevant, since it is associated with learning and behavioural disorders at school age (Hadders-Algra, 2002;Batstra *et al.*, 2003). Distinct neurological syndromes, such as cerebral palsy (CP), are classified as major neurological dysfunction (Hadders-Algra, 2003). We considered complex MND and major neurological dysfunction as an adverse neurological outcome.

In addition, the Hempel examination was used to calculate a neurological optimality score (NOS) and a fluency score by determining the number of items scored in a predefined optimality range (Huisman *et al.*, 1995). Subtle nervous system dysfunction is expressed in reduction of the fluency of movements; therefore the fluency score is sensitive for minimal changes in neuromotor development. Since the range for optimal behaviour is narrower than that for normal behaviour, the NOS is a valuable measure to evaluate subtle differences (Prechtl, 1980). Higher scores on the NOS and fluency scale represent better performance.

Parents were asked to complete the Dutch version of the Child Behaviour Check List (CBCL) 1½ - 5, a questionnaire designed and validated to identify problem behaviour in children aged 1½ to 5 years (Achenbach and Rescorla, 2000). CBCL items are grouped into the following problem scales; emotionally reactive, anxious/depressed, somatic complaints, withdrawn, sleep problems, attention problems and aggressive behaviour. The first four scales together form the internalizing scale (CBCL-int) and the latter two form the externalizing scale (CBCL-ext). The sum of all items determines the total problem scale (CBCL-tot). Higher scores represent more problematic behaviour.

Procedures

Demographic information, including e.g. parity, gestational age, birth weight, neonatal intensive-care unit admission, parental age, and educational level was collected on standardised charts at the follow-up assessments. At both centres, the assessors involved in the study were blinded to mode of conception.

Statistical analyses

We used Student's t-test and Fisher's exact test to compare demographic characteristics of the PGS and control groups. We used mixed-effects linear regression analysis to evaluate the possible effect of PGS on the outcome measures. Continuous outcome variables were transformed by the Box-Cox method in order

to achieve normality of regression residuals. The following transformations were used: (MDI/10)2.5, PDI, -ln(59.5-NOS), -ln(14.5-fluency), ln(CBCL-int), CBCL-ext, and ln(CBCL-tot). Apart from PGS we included the following fixed-effect variables in regression models: gender, gestational age (weeks), twin-status, maternal age at conception, hospital (where the assessments at the age of 2 years took place), birth weight, parity, prematurity status, conception via IVF or ICSI, age at examination and educational level of parents (low, medium or high). As data from the twins are likely to be correlated, the models included also a random effect due to mothers. We have used the results of the mixed-effects analysis to calculate confidence intervals (CI) for adjusted difference between the means of the PGS and control group. To interpret these intervals on the original scale we use the fact that the difference between means of two groups, A and B, on the transformed scale for the NOS score can be interpreted as the logarithm of the ratio (59.5-medB)/(59.5-medA), where medA and medB are medians on the original scale.

To evaluate the effect of PGS on the occurrence of an adverse neurological outcome we used the Fisher's exact test. A correction for potential confounders was not attempted and the dependence of twins was ignored because of a small number of non-normal outcomes.

Statistical analyses were performed using statistical packages R and SPSS 16. P-values of 5% or less were considered significant.

RESULTS

Between March 2004 and January 2006, the trial resulted in 52 ongoing pregnancies after PGS and 74 ongoing control pregnancies (pregnancies following conventional IVF-treatment) (Mastenbroek *et al.*, 2007). Respectively, 49 and 71 pregnancies resulted in the live birth of at least one child (table I). A pair of twins died directly postpartum due to immature birth and four couples withdrew informed consent during IVF-treatments and could therefore not be invited for follow-up. Reason for withdrawal was in most cases psychological stress caused by the blinding of couples to treatment allocation. Subsequently, 47 couples in the PGS group and 68 control couples were eligible for follow-up. Eventually, 45 (96%) PGS-couples with 54 children (36 singletons and 9 twins) and 63 (93%) control couples with 77 children (49 singletons and 14 twins) were included.

Demographic and neonatal characteristics of parents and children are shown in table II. Overall, characteristics of the groups were similar. Exceptions are that mean maternal age at conception was half a year higher in the control group

Table I - Eligibility and participation of couples and their children.

	Preimplantation genetic screening	Control group
Ongoing pregnancies	52	74
Intra-uterine deaths/ terminated pregnancies[a]	3	3
Pregnancies resulting in ≥ 1 live birth	49	71
Pregnancies resulting in postpartum death[b]	1	0
Withdrawal of informed consent[c]	1	3
Elegible couples	47	68
Non-participants[d]	2	5
Participating couples	45	63
with singletons	36	49
with twins	9	14
Children assessed at age 2 years[e]	54	77

[a] In the PGS-group: 1 termination because of trisomy 18, 1 intrauterine death due to abruptio placentae and 1 immature birth. In the control group: 2 terminations because of trisomy 18 and cleft lip and palate and 1 intrauterine death.
[b] postpartum death of twins due to immature birth at 24+5 weeks
[c] Couples whom withdrew informed consent during IVF-treatment were not approached for participation in the follow-up study.
[d] Reasons for non-participation: In the PGS group: 1 moving abroad, 1 assesment burden. In the control group: 1 moving abroad, 2 untraceable, 2 assessment burden.
[e] Analyses according to intention to treat

(P = .04) and gestational age was on average a week shorter in the control group (P = .03) compared to the PGS group.

Table III shows the unadjusted results of the neurological, psychomotor, mental and behavioural assessments. Mixed effects multiple regression analyses revealed no effect of PGS on MDI (P = .468), PDI (P = .540), fluency (P = .074), CBCL-int (P = .635), CBCL-ext (P = .583), and CBCL-tot (P = .442). A significant negative effect of PGS was found on NOS (P = .020). In the analyses the P-values are those from models including all 11 potential confounders. The best fitting models obtained by backward selection of variables included the following variables with a negative contribution to outcome: MDI, CBCL-int, CBCL-ext, and CBCL-tot: lower parental level of education; PDI: male gender and hospital (AMC), NOS: PGS, higher maternal age at conception, hospital (UMCG); and Fluency: higher maternal age at conception, hospital (UMCG). The estimated effect of PGS on -ln(59.5-NOS) was -0.219, the 95% CI was (-0.394 to -0.044). On the original NOS scale this result can be roughly interpreted as a difference of about 2 points (95% CI 0.5 to 4 points) between the corrected median of the PGS group and the corrected median of the control group. The raw unadjusted NOS medians of the two groups were 50 and 51, respectively.

Three children (6%) in the PGS group and one child (1%) in the control group were classified as having an adverse neurological outcome (P = .306, two-

sided). Three of these children had complex MND and one PGS-child was diagnosed with a diplegic cerebral palsy due to a rupture of an arteriovenous malformation at the age of 1 years.

Table II - Demographic characteristics of parents and children.

Characteristics	Couples with ≥ 1 live birth after PGS *(n = 45)*	Children born after PGS *(n = 54)*	Control-couples *(n = 63)*	Control-children *(n = 77)*	*P*-values
Parental characteristics:					
Maternal age at conception, mean (SD)	37.4 (1.4)		37.9 (1.6)		0.040
Primiparity	28 (62%)		46 (73%)		0.294
High education level mother[a]	23 (51%)		32 (51%)		1.000
High education level father[a]	25 (56%)		29 (46%)		0.435
Time to pregnancy in years, mean (SD)[b]	4.3 (2.3)		4.8 (2.9)		0.381
Cause of subfertility[c]:					
poor semen quality	19 (42%)		27 (43%)		
unexplained	13 (29%)		19 (30%)		
tubal	11 (24%)		17 (27%)		
anovulation	2 (4%)		3 (5%)		
endometriosis	2 (4%)		3 (5%)		
cervical	2 (4%)		2 (3%)		
ovarian failure[d]	1 (2%)		0		
Conception method:					0.995
IVF	26 (58%)		36 (57%)		
ICSI	16 (36%)		23 (37%)		
IUI[e]	1 (2%)		1 (2%)		
natural	2 (4%)		3 (5%)		
Perinatal characteristics:					
Twins	9 (20%)		14 (22%)		0.816
Gestational age in days, mean (SD)	275 (16)		268 (18)		0.027
Preterm birth (< 37 weeks)	5 (11%)		11 (18%)		0.420
Caesarean section[b]	8 (18%)		21 (33%)		0.121
Child characteristics:					
Male gender		27 (50%)		42 (54%)	0.722
Birth weight in grams, mean (SD)		3268 (711)		3028 (736)	0.064

[a] high level of education denotes university education or vocational colleges.
[b] 1 missing value.
[c] More than one diagnosis per couple was possible, statistics were not applied
[d] Donated oocytes of women of advanced maternal age were used.
[e] In case of poor follicle growth treatment was converted to intra-uterine insemination. This resulted in birth of a singleton in the PGS group and a twin in the control group.

DISCUSSION

The present study indicated that mental, psychomotor, and behavioural outcome at 2 years in children born following IVF with and without PGS was similar. Median scores on the tests in these domains were well within the normal range for both groups. Yet, PGS-children showed significantly lower neurological optimality scores than control children.

Since the range of optimal behaviour is narrower than the range of normal behaviour, the NOS and fluency score are sensitive measures for minor neurological dysfunction (Prechtl, 1980). The fact that we only found a statistically significant difference between the PGS and control group on the NOS and not on other psychomotor and neurological measures indicates that the neurodevelopmental difference between the two groups is rather subtle. The importance of this finding for long term neurological outcome is currently unknown and therefore continuing follow-up is important to warrant the safety of PGS.

The results of our study are mostly in accordance with the other neurodevelopmental follow-up studies performed so far. Banerjee *et al.* conducted a case-control study in which they compared 49 children born after embryo biopsy - PGS and preimplantation genetic diagnosis (PGD) – to 66 naturally conceived

Table III - Neurological outcome, behaviour and mental and motor perfomance.

Outcome measure	Children born after PGS *(n = 54)*	Control-children *(n = 77)*
Bayley Scales of Infant Development		
MDI, median (p25, p75)	103 (94.5, 110.0)[a]	103 (94.0, 110.5)
PDI, median (p25, p75)	92 (79.0, 103.0)[c]	90 (84.0, 99.0)
Neurological outcome Hempel		
Normal	47 (87%)	73 (95%)
Simple MND	4 (7%)	3 (4%)
Complex MND	2 (4%)	1 (1%)
Cerebral palsy	1 (2%)	0 (0%)
NOS, median (p25, p75)	50.0 (46.75, 52.0)	51.0 (47.0, 54.5)
fluency-score, median (p25, p75)	10.0 (9.0, 11.0)	11.0 (9.5, 11.0)
Child Behaviour Check List		
CBCL-int, median (p25, p75)	41.0 (37.0, 47.0)[a]	43.0 (37.0, 53.0)[b]
CBCL-ext, median (p25, p75)	47.0 (42.5, 56.0)[a]	51.0 (44.0, 56.0)[b]
CBCL-tot, median (p25, p75)	43.0 (38.0, 50.5)[a]	46.0 (41.0, 53.0)[b]

[a] 1 missing value
[b] 2 missing values
[c] 3 missing values

control children (Banerjee *et al.*, 2008). Remarkably, PGS/PGD-children scored significantly lower on the loco-motor subscale of the Griffiths than the controls, while the Griffiths General quotient and the results of the Toddler Temperament Questionnaire were similar in both groups. Nekkebroeck *et al.* compared a cohort of PGS/PGD-children with cohorts of ICSI and naturally conceived (NC) children (Nekkebroeck *et al.*, 2008a;Nekkebroeck *et al.*, 2008b). Consistent with our findings, they found no between-group differences for Bayley's MDI and PDI (Nekkebroeck *et al.*, 2008a). In addition, they found that PGS/PGD and ICSI mothers reported fewer CBCL total problems than NC mothers, whereas CBCL externalizing problems were reported fewer by ICSI mothers than PGS/PGD and NC mothers. ICSI fathers reported fewer total and external problems than PGS and NC fathers (Nekkebroeck *et al.*, 2008b).

Strengths of our study are its design, blinding of the assessors, minimal attrition, extensive correction for confounders, and the use of sensitive and validated neurodevelopmental outcome measures. This study is the first follow-up study of children born after IVF in which couples were randomised into IVF with or without PGS. Couples were informed on the follow-up when they were included in the trial (before start of the IVF-treatments), which resulted in high participation rates. Since twins as well as singletons were included, data are representative for children born following PGS.

A limitation of the present study is the fact that the power analysis was based on the number of women needed to detect an increase in ongoing pregnancy rate (Mastenbroek *et al.*, 2007) instead of the number of children needed for follow-up. The unforeseen negative effect of PGS on ongoing pregnancies further reduced the number of children available for follow-up. Future meta-analysis would be helpful to overcome sample size limitations of current follow-up studies. Uniformity of neurodevelopmental assessments is crucial for this matter.

Even though PGS, as practiced in the present trial, does not improve live birth rates, follow-up of children born after embryo-biopsy is still of importance. A recent meta-analysis confirmed that ongoing pregnancy rates following PGS are reduced (Mastenbroek *et al.*, 2008). Therefore the technique is no longer practiced on routine basis. Currently, it is studied whether alternative forms of PGS (such as polar body biopsy and more comprehensive chromosome testing) are feasible and efficient in improving live birth rates (Harper *et al.*, 2010;Geraedts *et al.*, 2010). Embryo-biopsy is also applied in PGD, a technique in which blastomeres are tested for specific heritable disorders in at-risk patients (Sermon *et al.*, 2004). Continuation of the use of embryo-biopsy justifies follow-up and reassessment of the children in the current trial at later ages may yield valuable information concerning the effects of embryo-biopsy on later development and health.

In conclusion, the results of the present follow-up study showed similar outcome in terms of mental, psychomotor and behavioural outcome in children born following PGS and conventional IVF. Nevertheless, PGS-children scored lower on the NOS, a marker for subtle neurological dysfunction. The importance of the latter finding for long term neurological outcome is currently unknown. Continuation of follow-up is needed to warrant the safety of PGS.