

Twins Born Following Fertility Treatment: Implications for Quantitative Genetic Studies

Adam Goody,¹ Frances Rice,^{1,2} Jacky Boivin,¹ Gordon T. Harold,¹ Dale F. Hay,¹ and Anita Thapar²

¹ The School of Psychology, Cardiff University, Cardiff, United Kingdom

² The Department of Psychological Medicine, School of Medicine, Cardiff University, Cardiff, United Kingdom

The rate of multiple births is substantially elevated in women who have had assisted reproduction treatment (ART; ~26%) compared to the general population (~1%), and these offspring are usually included in twin studies. Several studies have attempted to identify possible consequences of undergoing ART on the subsequent offspring. However, most studies have only included singleton births. We first examined whether twins born by ART differed from other twins on measures of childhood psychopathology, putative risk factors and correlates, and secondly tested for differences in the degree of twin similarity for available outcome measures. From a population-based twin sample, 101 families with dizygotic (DZ) twins conceived via ART were identified and compared with 1073 naturally conceived (NC) control DZ twin pairs. Analyses performed were (1) univariate and multivariate comparisons of between-group mean differences; and (2) comparison of twin 1–twin 2 correlations between the groups. The groups differed significantly on demographic factors (parental age, family size and social class) and pregnancy variables (smoking during pregnancy and birthweight) but did not differ on family conflict scores or in the frequency of obstetric complications. Family cohesion was higher in the ART group but this was accounted for by demographic factors. For child psychopathology there was a difference between the groups only for teacher-rated ADHD (Attention Deficit Hyperactivity Disorder). Differences were also found between groups for twin correlations. The differences found between ART and NC twins on group means and twin correlations suggest that researchers should be aware that including ART twins may influence results from twin studies.

In recent years there has been much interest in the development of children born from assisted reproduction treatment (ART). Approximately 1% to 3% of births in Europe occur via the use of some kind of fertility treatment (Nygren & Andersen, 2002), so it is of vital importance to examine the outcomes of children following ART. Rates of twin births are greatly increased following ART in comparison to naturally

conceived births. For example, in the United Kingdom (UK) approximately 26% of births following ART are multiple births, compared to only 1% of births following natural conception (Human Fertilisation and Embryology Authority, 1996). It is possible that ART is accounting for a greater proportion of twin births as ART becomes a more common method of conception. For instance, Bergh and colleagues (1999) examined rates of twin deliveries between 1986 and 1995 in a study of all births in Sweden. They found that in vitro fertilisation (IVF) pregnancies accounted for 2.2% of total twin deliveries between 1986 and 1990 but this figure rose to 12.9% between 1991 and 1995. The use of ART treatments and fertility drugs has been associated with an increase in rates of dizygotic (DZ) twinning (Hall, 2003; Macdonald, 1993).

Several studies have been undertaken to investigate the characteristics of ART families and outcomes for children born following the use of these techniques. The majority of research has found that ART families are comparable to families of naturally conceived (NC) children in terms of parental anxiety and wellbeing during pregnancy (Colpin et al., 1999; McMahan et al., 1995; McMahan et al., 1997), and subsequent child psychological outcomes (Golombok et al., 1996; Golombok et al., 2002; Golombok et al., 1995). Some studies have reported more positive outcomes for parents who have children following ART such as lower parenting stress and higher levels of warmth and emotional involvement with their children (Golombok et al., 1996; Golombok et al., 1995). A small number of studies have examined psychological characteristics and outcomes for twins born following ART (Cook et al., 1998; Tully et al., 2003). However, with the exception of one study, these studies have been of very small samples. Tully and colleagues (2003) examined the psychological characteristics and outcomes of families of 5-year-old twins

Received 15 March, 2005; accepted 21 April, 2005.

Address for correspondence: Frances Rice, Department of Psychological Medicine, School of Medicine, Cardiff University, Heath Park, Cardiff CF14 4XN, UK. E-mail: Ricef@cardiff.ac.uk

born from ART. This study used a matched control design, with 121 twin pairs in each group (ART vs. NC), and found no differences on a range of parental adjustment and child behavior outcomes rated by both parent and teacher, as well as the child's birthweight. The only significant finding was that parents of NC children were more inconsistent in their use of discipline. Apart from this study, the vast majority of studies of children born by ART have included only singletons.

Higher rates of obstetric and perinatal complications for ART mothers and their children have also been reported, including preterm delivery, low birthweight and neurological sequelae such as cerebral palsy and developmental delay (Bergh et al., 1999; Stromberg et al., 2002). However, studies of obstetric and perinatal factors that have focused on twins born via ART have suggested that differences found between ART and NC groups are reduced or even eliminated when only groups of twins are compared. For example, Pinborg and colleagues, (Pinborg, Loft, Rasmussen, et al., 2004; Pinborg, Loft, Schmidt, et al., 2004) compared twins born from IVF or intracytoplasmic sperm injection (ICSI) with NC twins and found no differences in birthweight, gestational age, and frequency of Caesarean sections after controlling for maternal age and parity. Similarly, when multiple birth was controlled for, no differences in the incidence of neurological sequelae were found in the study of Stromberg and colleagues (2002). There has, however, been relatively consistent evidence that children conceived via ART have lower birthweights than naturally conceived children even in singletons (Schieve et al., 2002).

Finally, there has been recent evidence suggesting that children born by ART may show increased rates of specific medical disorders involving epigenetic alterations (Niemitz & Feinberg, 2004). In particular, elevated rates of disorders that involve imprinting, such as Beckwith Weiderman Syndrome and Angelman Syndrome have been reported. This raises the question that levels of psychopathology or twin similarity may vary due to elevated rates of imprinting in twins born by ART. Epigenetic effects are assumed to be stochastic, therefore, one would assume that an increase in epigenetic alterations would decrease similarity between twins conceived via ART. That is, the early preimplantation that twins share would be less similar in twins conceived via ART compared to NC twins.

As the rate of twin births following fertility treatment increases in many countries, it is important to assess the influence that this increase may have on behavioral genetic studies. First, an increase in the number of twins born via ART will primarily serve to increase the number of DZ twin pairs in a twin sample (Hall, 2003). The effect this would have on genetic and environmental estimates from behavioral genetic studies is unclear. Second, Stoolmiller (1999)

has argued that genetic and environmental estimates derived from adoption studies may be influenced by restricted range of family environments resulting from only families at low-environmental risk being able to adopt children. Thus, it is possible that including ART twins in a twin study may influence genetic and environmental parameter estimates if ART twins are not comparable to naturally conceived twins in terms of environmental risk. Third, if preimplantation environment and imprinting has important, wide-ranging effects on outcomes (medical outcomes and psychopathology), one might expect decreased twin similarity for outcomes in ART compared to twins who are conceived naturally.

The present study used a population-based twin sample to assess (1) whether twins born by ART differ from those born by NC on key demographic, pregnancy and obstetric variables, as well as childhood psychopathology and educational difficulties; and (2) whether twin correlations differ between the ART and NC groups on childhood psychopathology and birthweight. We examined twin similarity on birthweight given the reported mean difference in birthweight between children conceived via ART and NC children and because imprinted genes are involved in determining birthweight.

Method

Sample

Data were collected by postal questionnaire in 1996 as part of the Cardiff Study of All Wales and North-West England Twins (CASTANET). The population-based sample of school-age twin pairs had been identified from nine health districts in Greater Manchester and Lancashire, UK. Of 2846 families contacted, 2082 completed and returned a package of questionnaires (response rate = 73%). Of these, 1607 families gave permission to contact teachers in order to obtain teacher-rated data (77%), and of the teachers, 1470 replied (response rate = 91.5%).

Measures

Pregnancy and Obstetrics

Mothers reported on pregnancy and obstetric complications using an adapted version of the Lewis and Murray (1987) scale. Except in the case of smoking during pregnancy, we had no a priori hypotheses that groups should differ on specific pregnancy and obstetric risk variables. Therefore, to avoid running several tests on binary variables, composite measures of 'pregnancy risk' and 'obstetric risk' were computed. Pregnancy risk included (1) number of cigarettes smoked during pregnancy (0, less than 10, between 10 and 20, greater than 20 cigarettes per day); (2) admission to hospital because of high blood pressure or swollen hands/feet; and (3) vaginal bleeding during pregnancy. A pregnancy risk score was obtained by summing these three items. The range of scores was 0 to 5, with higher scores indicating

increased risk. A variable 'delivery risk' was computed and included (1) having an emergency Caesarean section; (2) a forceps or ventouse delivery; and (3) a labour shorter than 3 hours or longer than 36 hours. These items were summed to obtain a delivery risk score ranging from 0 to 3, a higher score indicating increased risk.

As prevalence of smoking is higher among lower social classes, and previous studies have suggested that ART families may differ from NC families in terms of social class (Golombok et al., 2001), smoking was also coded as a binary variable (presence/absence of smoking during pregnancy) and was analyzed separately. Other obstetric risk variables measured were gestation period in weeks, birthweight in grams, and whether the twins were kept in a Special Care Baby Unit (SCBU) or incubator.

Child Psychopathology

Attention Deficit Hyperactivity Disorder (ADHD)

Symptoms of child psychopathology were reported by mothers and by teachers where consent was given. A modified version of the DuPaul ADHD rating scale (DuPaul, 1991; Thapar et al., 2000) was used to assess ADHD symptoms. This measure included the original 14 *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed., rev.; DSM-III-R, American Psychiatric Association, 1987) symptoms of ADHD plus four items to measure the additional symptoms of ADHD and Hyperkinetic disorder found in ICD-10 (World Health Organization, 1993) and DSM-IV (American Psychiatric Association, 1994) criteria. The items were summed to obtain a total ADHD score ranging from 0 to 54, with a higher score indicating a higher presence of ADHD symptoms (Cronbach's $\alpha = .96$ for mother-rated data, $\alpha = .97$ for teacher-rated data).

Internalizing Symptoms

Internalizing symptoms were assessed using six items from the Rutter A scale for mother-rated data and six items from the Rutter B scale for teacher-rated data (Rutter et al., 1970). These items were 'often worried, worries about many things', 'tends to be fearful or afraid of new things or situations', 'often appears miserable, unhappy, tearful or distressed', 'tends to be on his/her own, rather solitary', 'any sleeping difficulties?', 'any eating difficulties?' and were scored on a three point scale (0 to 2). Scores range from 0 to 12 for both scales, with higher scores indicating a higher presence of internalizing symptoms. Cronbach's $\alpha = .65$ for the Rutter A scale and $.80$ for the Rutter B scale.

Antisocial Behavior

Antisocial behavior was assessed using the 'conduct difficulties' subscale of the Rutter A and B scales. The Rutter A conduct difficulties scale consists of five items which are summed to obtain a total score ranging from 0 to 10 ($\alpha = .81$). The Rutter B conduct difficulties scale

consists of six items, with a total score range of 0 to 12 ($\alpha = .89$). For both scales a higher score indicates a higher level of antisocial behavior.

Family Environment

Family conflict and family cohesion was measured by maternal report of the Family Environment Scale (Moos & Moos, 1981). The conflict and cohesion subscales consist of nine items with a true/false response format, giving a possible range of scores from 0 to 9 for each scale. Higher scores indicate higher levels of conflict/cohesion respectively in the family environment ($\alpha = .73$ for conflict and $.67$ for cohesion).

Educational Difficulties

Mothers were also asked to report on certain educational difficulties: (1) whether the child was currently behind in reading; and (2) whether they had a learning difficulty.

Data Analysis

As a result of analyzing twin pairs, some of the variables did not contain independent observations as each twin pair provided two data points. For univariate analysis, nonindependence was corrected for via the survey command functions in Stata 6.0 (StataCorp, 1999). This is done by likening the twin data to a two-stage cluster design with the twin pairs as the primary sampling unit, and adjusting the variances of the analyses to be equivalent to that of independently sampled pairs. For multivariate analysis, this option is not available in Stata 6.0 (StataCorp, 1999). Therefore, a random sample of twins was taken to ensure that: (a) an equal number of first born and second born twins were selected, and (b) that only one twin was selected from each twin pair.

Categorical variables were analyzed using Pearson's χ^2 test, while MANOVAs were run on the remaining continuous variables for family demographics, family environment, obstetric complications, parent-rated child psychopathology and teacher-rated psychopathology. SPSS version 11.0 (2001) was used for analysis of independent data, while Stata 6.0 (StataCorp, 1999) was used for χ^2 tests on nonindependent data. Pearson's r was used to analyze correlations between twin pairs on child psychopathology and birthweight, using SPSS version 11.0 (2001).

Results

Sample Characteristics

Out of the original total sample, 1920 twin pairs could be assigned zygosity. Zygosity was determined using a twin similarity questionnaire and an algorithm based on previous work (Cohen et al., 1975; Thapar & McGuffin, 1994). Accuracy of greater than 90% in assigning zygosity has been reported (Thapar et al., 2000). Of the DZ twins, 101 pairs (8.5% of the total sample of DZ twins) were identified where the

parents reported having used some type of fertility treatment, which formed the ART group. Fifteen twin pairs were excluded from the sample as the parents had not indicated whether they had used fertility treatment or not. The remaining DZ twin pairs who reported not using fertility treatment formed the naturally conceiving (NC) group ($n = 1073$). Parent-rated data were available for all of these twin pairs. Teacher-rated data were available for 743 NC twin pairs and 67 ART twin pairs. Of the sample involved in the current report, 48.2% of children were boys and 51.8% were girls. The mean maternal age at birth was 28.56 (range 16 to 57) and mean paternal age at birth was 31.23 (range 17 to 63). The average age of the twins in the study was 11.12 (range 5 to 17). The distribution of age for the twins from each group was similar although children from the ART group were on average 6 months younger (mean = 10.54, range 6 to 17) than children in the NC group (mean = 11.16, range 5 to 17; $t = -1.902$, $df = 1170$, $p = .057$). The majority of participants described their ethnicity as British (1083 mothers, 1069 fathers). The remainder described themselves as other European (11 mothers, 13 fathers), African/Caribbean (13 mothers, 17 fathers), Bangladeshi/Indian/Pakistani (41 mothers, 44 fathers), Jewish (10 mothers, 10 fathers), Arab (2 mothers, 4 fathers), South-East Asian (2 mothers, 2 fathers), and other (7 mothers, 8 fathers). The ethnic background of parents in the ART and the NC groups were broadly comparable, for example, the proportion of mothers describing their ethnicity as British was 93% in both the ART and the NC groups. Although small numbers in some categories made it difficult to meaningfully interpret percentages, there was a slightly greater proportion of parents in the NC group who described themselves as Bangladeshi, Indian or Pakistani (ART mothers 1%, ART fathers 2%, NC mothers 3.7%, NC fathers 3.9%). Social class was classified according to Standard Occupational Classification Vol. 1 (Office for

National Statistics, 1995). The proportions in each category were: professional 12.0%, managerial/technical 29.5%, skilled occupations 28.4%, semiskilled 4.7%, unskilled 1.8%, unemployed/sick/disabled or studying 22.1%. Families where completed replies were also received from teachers were comparable to the whole sample in terms of social class ($\phi = .025$, $p = .994$) and employment status of main earner ($\phi = .043$, $p = .904$).

Family Demographics

Table 1 shows a significant multivariate group difference for demographic variables, $\lambda = .954$, $F(4, 1115) = 13.547$, $p = .001$. The groups differed significantly by maternal age at birth, $F = 5.22$, $p = .023$, paternal age at birth, $F = 4.73$, $p = .030$, the number of the twin's siblings, $F = 36.31$, $p = .001$, and social class, $F = 15.98$, $p = .001$. Twins in the ART group tended to belong to a higher social class, had significantly older parents, and had significantly fewer siblings than twins in the NC group.

Family Environment

There was no significant multivariate group difference on the family environment variables, $\lambda = .996$, $F(2, 1064) = 2.20$, $p = .112$. Neither variable differed significantly by group, however there was a trend, $p = .06$ for higher cohesion scores in the ART group (Table 1). Although the difference in cohesion scores did not reach significance according to conventional levels, an analysis of covariance was performed controlling for family demographic factors shown in Table 1 in order to test whether these demographic factors accounted for the group difference in family cohesion. There was no group difference on family cohesion when demographic factors were included as covariates. A higher maternal age at childbirth, $F = 4.483$, $p = .034$, and a lower number of siblings, $F = 5.945$, $p = .015$, were significant covariates associated with increased levels of family cohesion.

Table 1

Family Demographics and Environment

Family demographics

$\lambda = .954$, $F = 13.547$, $p = .001$

Variable	<i>N</i>	NC mean	ART mean	<i>F</i> value	<i>p</i> value
Mother's age at birth	1157	28.43	29.61	5.22	.023
Father's age at birth	1141	31.11	32.51	4.73	.030
Number of siblings	1167	1.64	0.80	36.31	.001
Social class	1157	3.80	2.97	15.98	.001

Family environment scale

$\lambda = .996$, $F = 2.20$, $p = .112$

Variable	<i>n</i>	NC mean	ART mean	<i>F</i> value	<i>p</i> value
Cohesion	1094	7.42	7.73	3.57	.059
Conflict	1126	3.15	3.17	0.001	.977

Table 2

Pregnancy, Obstetric and Educational Outcomes

Pregnancy and obstetric variables

Lambda = .997, $F = 1.089$, $p = .353$

	<i>n</i>	NC mean	ART mean	<i>F</i> value	<i>p</i> value
Pregnancy risk	1153	1.04	.99	0.20	.641
Delivery risk	1041	.86	.98	2.94	.088
Gestation period (weeks)	1129	36.35	36.44	0.03	.844
	<i>n</i>	NC mean	ART mean	<i>t</i> value	<i>p</i> value
Birthweight (grams)	1166	2473.87	2284.53	-2.642	.008
Binary pregnancy and obstetric variables	<i>n</i>	'yes' frequency NC group	'yes' frequency ART group	χ^2	<i>p</i> value
Smoking during pregnancy	1171	31.1%	17.8%	7.78	.005
SCBU	2344	37.2%	35.1%	0.32	.673
Incubator	2343	34.0%	33.8%	0.002	.973
Educational difficulties					
Behind in reading	2344	18.7%	14.9%	1.84	.189
Learning difficulty	2339	13.7%	11.4%	0.82	.402

Obstetric Complications

Table 2 showed no significant multivariate group difference for the continuous obstetric outcomes, lambda = .997, $F(3, 1029) = 1.089$, $p = .353$. The only obstetric complication to approach significance was delivery risk, $F = 2.942$, $p = .09$, with higher scores in the ART group.

Each twin provided a unique data point for birthweight, therefore, birthweight was analyzed separately with correction for nonindependence as described in the methods section. Twins in the ART group were significantly lighter at birth than those in the NC group, $t = -2.642$, $p = .008$. The birthweight of twins in the ART group remained lower than that of the NC group after adjusting for demographic factors, smoking during pregnancy and length of gestation, $F = 11.589$, $p = .001$.

For binary obstetric variables, the frequency of mothers who smoked during pregnancy was significantly higher in the NC group compared with the ART group, $\chi^2 = 7.78$, $df = 1$, $p = .005$. There were no differences between groups on admission to a SCBU or incubator.

Educational Difficulties

Neither of the variables on educational difficulties differed significantly between the two groups (Table 2).

Childhood Psychopathology: Parent-Rated

Table 3 shows that there was no significant multivariate group difference for parent-rated psychopathology variables, lambda = .998, $F(3, 1136) = .846$, $p = .469$. None of the variables in the model differed significantly by group.

Childhood Psychopathology: Teacher-Rated

Table 3 shows no significant group multivariate difference for the three teacher-rated psychopathology variables in the model, lambda = .990, $F(3, 765) = 2.480$, $p = .060$. However, teacher-rated ADHD was found to differ significantly between groups, $F = 6.87$, $p = .009$, with a higher mean ADHD score in the NC group. Given this result, an ANCOVA that included the demographic variables showing significant group differences (see Table 1) as well as maternal smoking during pregnancy as covariates was performed. When these factors were included as covariates, there was a trend for a group difference, $F = 3.685$, $p = .055$. Social class was the only significant covariate, $F = 12.226$, $p = .001$.

Twin Correlations

Table 4 shows that there were significant multivariate differences on twin correlations for birthweight (Fisher's $z = -2.01$, $p = .023$), parent-rated internalizing ($z = -1.99$, $p = .023$) and antisocial behavior ($z = -2.22$, $p = .014$). For all three of these variables, twin correlations in the ART group were lower than for the NC group.

Discussion

The present study utilized a population-based sample, thus eliminating cohort effects which may occur in studies selecting ART families from fertility clinics and matching them to a control group. The results indicate that there are differences between the ART and NC groups on a range of variables. Of the demographic variables measured, ART families were found to have a higher maternal age, paternal age and social class. A higher maternal and paternal age has been demon-

Table 3

Childhood Psychopathology

Childhood psychopathology parent-rated

Lambda = .998, $F = .846$, $p = .469$

	<i>n</i>	NC mean	ART mean	<i>F</i> value	<i>p</i> value
Internalizing	1147	1.71	1.63	0.02	.881
ADHD	1162	12.09	11.69	0.07	.794
Antisocial behavior	1173	1.24	1.05	1.26	.262

Child psychopathology teacher-rated

Lambda = .990, $F = 2.480$, $p = .060$

	<i>n</i>	NC mean	ART mean	<i>F</i> value	<i>p</i> value
Internalizing	783	1.40	1.19	0.76	.382
ADHD	791	10.46	6.23	6.87	.009
Antisocial behavior	783	.96	.69	1.31	.254

strated in previous studies of IVF singletons (Bergh et al., 1999; McMahan et al., 2003). However, this was not found in the only previous large study of IVF twins (Tully et al., 2003). It is possible that the consistent finding in singleton studies will be reduced in a twin sample as fraternal twinning rates increase with an increase in maternal age (Astolfi et al., 2003); therefore mothers (and fathers) of naturally conceiving twins will tend to be older than mothers (and fathers) of singletons. A lower number of siblings among IVF families has also been demonstrated previously (Bergh et al., 1999) and is to be expected as families having fertility treatment are likely to have decreased fecundity and, in addition, have reduced opportunities for further conception because the first offspring is conceived relatively late in the fertility life cycle.

No significant differences were found between groups on the family environment variables of cohesion and conflict although there was a trend for increased cohesion in the ART group. This trend for increased cohesion in the ART group was accounted for demographic features, in particular, a higher maternal age and a lower number of other children in the ART group. These results are in line with those of Tully and colleagues (2003) who found no differences on measures which assessed the quality of the mother-partner relationship and with a number of studies that have reported slight increases in parental competence and warmth in families who have used ART (van Balen, 1998). Tully and colleagues matched control (NC) families to ART families on a range of potentially confounding variables including family income, child birthweight and maternal age and tested for group differences on child psychopathology and parenting. For the purposes of the present study, however, it was important to test whether there were differences between the groups on demographic variables as this may impact on the representativeness of a twin sample.

The differences found on demographic variables suggest that ART twin families represent a low-risk group compared with NC twin families. It is therefore likely that as the number of twins born following fertility treatment increases, the proportion of twins exposed to high-risk environments will correspondingly decrease. As mentioned previously, a restriction of range such as this may influence genetic and environmental parameter estimates from twin studies. Nevertheless, there is an increasing focus on the transfer of single embryos in Europe (Saldeen & Sundstrom, 2005). This suggests that, at least in some countries, the increased rates of multiple births currently seen in ART (World Health Organization, 2002) may diminish in the future.

Having found differences on a range of demographic variables between the groups, a further question was whether this would be associated with obstetric and child psychopathology variables. Demographic variables have been associated with obstetric outcomes (Sameroff & Chandler, 1975) and childhood psychopathology (Rutter & Sroufe, 2000). In the present study, smoking in pregnancy when ana-

Table 4

DZ Twin Correlations by Group for Birthweight and Psychopathology Variables

	NC twin <i>r</i>	ART twin <i>r</i>	Fisher's <i>z</i>	<i>p</i> value
Birthweight	.711	.588	-2.01	.023
Parent-rated:				
Internalizing	.275	.068	-1.99	.023
ADHD	.324	.316	-.09	.500
Antisocial behavior	.372	.274	-1.04	.147
Teacher-rated:				
Internalizing	.309	.362	0.395	.345
ADHD	.431	.279	-1.33	.089
Antisocial behavior	.459	.203	-2.22	.014

lyzed as a binary variable (any smoking during pregnancy vs. no smoking during pregnancy) showed a significant difference with 38% of NC mothers smoking versus 18% of ART mothers. This is consistent with previous work that has found that smoking is correlated with social class (Shewry et al., 1992). However, it is also plausible that the ART group, who highly desire a child and will have planned for pregnancy, will be more likely to abstain from smoking during pregnancy. Without data on rates of smoking before the pregnancy, this possibility cannot be tested. Mean birthweight was significantly lower for the ART group than for the NC group even after adjustment for potential confounders, and similar findings have previously been reported by other studies (Lambalk & van Hooff, 2001; Pinborg, Loft, Schmidt, et al., 2004). Therefore, while the ART twins in the present sample are a lower risk group in terms of demographic variables and rates of smoking during pregnancy, the twin birthweights of the ART groups are lower, which may impact on later health and behavior (Barker, 1998; Bhutta et al., 2002).

The multivariate models of parent and teacher-rated psychopathology showed no significant main effect group differences on psychopathology variables, although the teacher-rated model was close to achieving significance. Only teacher-rated ADHD scores differed significantly between groups, with the ART group reporting lower ADHD scores. However, no group difference was found for parent-rated ADHD scores. Nevertheless, consistent with the findings on demographic variables this finding for teacher-rated ADHD may indicate the presence of a lower risk group.

The final analysis involved comparing the two groups by twin 1–twin 2 correlations. Twin correlations were significantly lower for the ART group on birthweight, parent-rated internalizing and teacher-rated antisocial behavior. It is also worth noting that except for teacher-rated internalizing, the trend for all other variables was for ART twins to have lower twin correlations than NC twins. The observed differences in twin correlations could clearly have an effect on genetic and environmental parameter estimates drawn from twin studies and suggest that ART is a potential source of heterogeneity in twin analysis and should be considered when analyzing data. It is possible that decreased correlations for ART twins may be due to greater differences in the preimplantation environment for ART compared to NC twins.

As mentioned previously, a higher proportion of DZ twins will be born via ART compared to MZ twins, so in any given twin sample one would expect a higher proportion of ART families in the DZ twin group. The results of this study suggest that the increase in numbers of DZ twins born via fertility treatment may influence genetic and environmental parameter estimates from twin studies. First, the differences found on demographic as well as one prenatal

variable (maternal smoking in pregnancy) between the groups will act to restrict the range of environments in DZ twin families, and restricted range has been shown previously to affect parameter estimates within genetic designs (Stoolmiller, 1999). However, the twins born by ART are not just uniformly 'lower risk', as they showed lower birthweight. Second, differences were found between the groups on DZ twin correlations suggesting that increasing rates of twins born following ART may impact upon genetic and environmental parameter estimates derived from twin studies. The lower rates of teacher-rated ADHD further suggest that the differences between groups on demographic variables and rates of smoking during pregnancy may be impacting upon outcome variables. However, this is not conclusive as this result was not found for parent-rated ADHD. Collectively, the results presented here suggest that twin researchers should be aware of possible effects of the increase in DZ twins born via ART, and the importance of obtaining information on fertility treatment in the process of data collection and testing for heterogeneity in the twin sample.

Acknowledgments

The twin register was set up with funding from the Medical Research Council, UK.

References

- American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev.). Washington, DC: Author.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- Astolfi, P., Ulizzi, L., & Zonta, L. A. (2003). Changes in twinning rate: Italy 1950–1996. *Human Reproduction*, *18*, 207–211.
- Barker, D. J. P. (1998). *Mothers, babies and health in later life* (2nd ed.). Edinburgh, UK: Churchill Livingstone.
- Bergh, T., Ericson, A., Hillensjo, T., Nygren, K. G., & Wennerholm, U. B. (1999). Deliveries and children born after in-vitro fertilisation in Sweden 1982–95: A retrospective cohort study. *Lancet*, *354*, 1579–1585.
- Bhutta, A. T., Cleves, M. A., Casey, P. H., Cradock, M. M., & Arnand, K. J. S. (2002). Cognitive and behavioral outcomes of school-aged children who were born preterm: A meta-analysis. *Journal of the American Medical Association*, *288*, 728–737.
- Cohen, D. J., Dibble, E., Grawe, J. M., & Pollin, W. (1975). Reliably separating identical from fraternal twins. *Archives of General Psychiatry*, *32*, 1371–1375.
- Colpin, H., Munter, A. D., Nys, K., & Vandemeulebroecke, L. (1999). Parenting stress and psychosocial well-being among parents with twins conceived naturally or by reproductive technology. *Human Reproduction*, *14*, 3133–3137.

- Cook, R., Bradley, S., & Golombok, S. (1998). A preliminary study of parental stress and child behavior in families with twins conceived by in-vitro fertilization. *Human Reproduction*, *13*, 3244–3246.
- DuPaul, G. J. (1991). Parent and teacher ratings of ADHD symptoms: Psychometric properties in a community-based sample. *Journal of Clinical Child Psychology*, *20*, 245–253.
- Golombok, S., Brewaeys, A., Cook, R., Giavazzi, M. T., Guerra, D., Mantovani, A., van Hall, E., Crosignani, P. G., & Dexeus, S. (1996). The European study of assisted reproduction families: Family functioning and child development. *Human Reproduction*, *11*, 2324–2331.
- Golombok, S., Brewaeys, A., Giavazzi, M. T., Guerra, D., MacCallum, F., & Rust, J. (2002). The European study of assisted reproduction families: The transition to adolescence. *Human Reproduction*, *17*, 830–840.
- Golombok, S., Cook, R., Bish, A., & Murray, C. (1995). Families created by the new reproductive technologies: Quality of parenting and social and emotional development of the children. *Child Development*, *66*, 285–298.
- Golombok, S., MacCallum, F., & Goodman, E. (2001). The 'test-tube' generation: Parent-child relationships and the psychological well-being of in vitro fertilization children at adolescence. *Child Development*, *72*, 599–608.
- Golombok, S., Murray, C., Brinsden, P., & Abdalla, H. (1999). Social versus biological parenting: Family functioning and the socioemotional development of children conceived by egg or sperm donation. *Journal of Child Psychology and Psychiatry*, *40*, 519–527.
- Hall, J. G. (2003). Twinning. *Lancet*, *362*, 735–743.
- Human Fertilisation and Embryology Authority. (1996). *Fifth Annual Report: July 1996*. Annex 3.
- Koivurova, S., Hartikainen, A. L., Sovio, U., Gissler, M., Hemminki, E., & Jarvelin, M. R. (2003). Growth, psychomotor development and morbidity up to 3 years of age in children born after IVF. *Human Reproduction*, *18*, 2328–2336.
- Lambalk, C. B., & van Hooff, M. (2001). Natural versus induced twinning and pregnancy outcome: A Dutch nationwide survey of primiparous dizygotic twin deliveries. *Fertility and Sterility*, *75*, 731–736.
- Lewis, S. W., & Murray, R. M. (1987). Obstetric complications, neurodevelopmental deviance, and risk of schizophrenia. *Journal of Psychiatric Research*, *21*, 413–421.
- Macdonald, A. M. (1993). Twin studies in medical research. *Lancet*, *341*, 1419.
- McMahon, C., Gibson, F., Leslie, G., Cohen, J., & Tennant, C. (2003). Parents of 5-year-old in vitro fertilization children: Psychological adjustment, parenting stress, and the influence of subsequent in vitro fertilization treatment. *Journal of Family Psychology*, *17*, 361–369.
- McMahon, C. A., Ungerer, J. A., Beaupaire, J., & Tennant, C. (1995). Psychosocial outcomes for parents and children after in vitro fertilization: A review. *Journal of Reproductive and Infant Psychology*, *13*, 1–16.
- McMahon, C. A., Ungerer, J. A., Beaurepaire, J., Tennant, C., & Saunders, D. (1997). Anxiety during pregnancy and fetal attachment after in-vitro fertilization conception. *Human Reproduction*, *12*, 176–182.
- Moos, R. H., & Moos, B. H. (1981). *Manual for the family environment scale*. Palo Alto, CA: Consulting Psychologists Press.
- Niemitz, E. L., & Feinberg, A. P. (2004). Epigenetics and assisted reproductive technology: A call for investigation. *American Journal of Human Genetics*, *74*, 599–609.
- Nyboe Andersen, A., Gianaroli, L., & Nygren, K. G. (2004). Assisted reproductive technology in Europe, 2000: Results generated from European registers by ESHRE. *Human Reproduction*, *19*, 490–503.
- Nygren, K. G., & Andersen, A. N. (2002). Assisted reproductive technology in Europe, 1999: Results generated from European registers by ESHRE. *Human Reproduction*, *17*, 3260–3274.
- Office for National Statistics. (1995). *Standard occupational classification* (Vol. 2). London: The Stationary Office.
- Pinborg, A., Loft, A., Rasmussen, S., Schmidt, L., Langhoff-Roos, J., Greisen, G., & Andersen, A.N. (2004a). Neonatal outcome in a Danish national cohort of 3438 IVF/ICSI and 10,362 non-IVF/ICSI twins born between 1995 and 2000. *Human Reproduction*, *19*, 435–441.
- Pinborg, A., Loft, A., Schmidt, L., Langhoff-Roos, J., & Andersen, A. N. (2004b). Maternal risks and perinatal outcome in a Danish national cohort of 1005 twin pregnancies: The role of in vitro fertilization. *Acta Obstetrica et Gynecologica Scandinavica*, *83*, 75–84.
- Rutter, M., & Sroufe, L. A. (2000). Developmental psychopathology: Concepts and challenges. *Development and Psychopathology*, *12*, 265–296.
- Rutter, M., Tizard, J., & Whitmore, K. (1970). *Education, health and behavior*. London: Longman.
- Saldeen, P., & Sundstrom, P. (2005). Would legislation imposing single embryo transfer be a feasible way to reduce the rate of multiple pregnancies after IVF treatment? *Human Reproduction*, *20*, 4–8.
- Sameroff, A. J., & Chandler, M. J. (1975). Reproductive risk and the continuum of caretaking casualty. In F. D. Horowitz (Ed.), *Review of child development research* (Vol. 4; pp. 187–244). Chicago: University of Chicago Press.
- Schieve, L. A., Meikle, S. F., Ferre, C., Peterson, H. B., Jeng, G., & Wilcox, L. S. (2002). Low and very low birthweight in infants conceived with use of assorted reproductive technology. *New England Journal of Medicine*, *346*, 731–737.

- Shewry, M. C., Smith, W. C., Woodward, M., & Tunstall-Pedoe, H. (1992). Variation in coronary risk factors by social status: Results from the Scottish Heart Health Study. *British Journal of General Practice*, 42, 406–410.
- SPSS Inc. (2001). SPSS for Windows (Release 11.0.1) [Computer software]. Chicago: SPSS, Inc.
- StataCorp. (1999). Stata statistical software (Release 6.0) [Computer software]. College Station, TX: Stata Corporation.
- Stoolmiller, M. (1999). Implications of the restricted range of family environments for estimates of heritability and nonshared environment in behavior-genetic adoption studies. *Psychological Bulletin*, 125, 392–409.
- Stromberg, B., Dahlquist, G., Ericson, A., Finnstrom, O., Koster, M., & Stjernqvist, K. (2002). Neurological sequelae in children born after in-vitro fertilisation: A population-based study. *Lancet*, 359, 461–465.
- Thapar, A., Harrington, R., Ross, K., & McGuffin, P. (2000). Does the definition of ADHD affect heritability? *Journal of the American Academy of Child and Adolescent Psychiatry*, 39, 1528–1536.
- Thapar, A., & McGuffin, P. (1994). A twin study of depressive symptoms in childhood. *British Journal of Psychiatry*, 165, 259–265.
- Tully, L. A., Moffitt, T. B., & Caspi, A. (2003). Maternal adjustment, parenting and child behavior in families of school-aged twins conceived after IVF and ovulation induction. *Journal of Child Psychology and Psychiatry*, 44, 316–325.
- van Balen, F. (1998). Development of IVF children. *Developmental Review*, 18, 30–46.
- Westergaard, T., Wohlfahrt, J., Aaby, P., & Melbye, M. (1997). Population based study of rates of multiple pregnancies in Denmark, 1980–94. *British Medical Journal*, 314, 775–779.
- World Health Organization. (1993). *The ICD-10 classification of mental and behavioural disorders: Diagnostic criteria for research*. Geneva, Switzerland: Author.
- World Health Organization (2002). *Current practices and controversies in assisted reproduction*. In E. Vayena, P. J. Rowe, & P. D. Griffin (Eds.), Report of a meeting on ‘Medical, Ethical and Social Aspects of Assisted Reproduction’ held at WHO Headquarters in Geneva, Switzerland 17–21 September, 2001.
-