International Collaborative Study of Intracytoplasmic Sperm Injection–Conceived, In Vitro Fertilization–Conceived, and Naturally Conceived 5-Year-Old Child Outcomes: Cognitive and Motor Assessments

I. Ponjaert-Kristoffersen, PhD*; M. Bonduelle, MD, PhD‡; J. Barnes, PhD§; J. Nekkebroeck, PhD*; A. Loft, MD‖; U.-B. Wennerholm, MD, PhD¶; B. C. Tarlatzis, MD#; C. Peters, MD**; B. S. Hagberg, PhD¶¶; A. Berner, PhD††; and A. G. Sutcliffe, MD**

ABSTRACT. Objective. To date, very few studies have been conducted on the neurodevelopmental wellbeing of children conceived through intracytoplasmic sperm injection (ICSI). The limitations of these studies often include a lack of comparison with a demographically matched, naturally conceived (NC) group and the investigation of only very young children, with relatively small samples sizes. One study showed that there were no differences in IQ scores among ICSI-conceived, in vitro fertilization (IVF)-conceived, and NC children at 5 years of age. Unfortunately, psychomotor development was not assessed in that study. Because findings regarding these children’s cognitive and motor development are inconclusive, the aim of this study was to shed more light on the cognitive and motor development of 5-year-old ICSI-conceived children.

Methods. A total of 511 ICSI-conceived children were compared with 424 IVF-conceived children and 488 NC controls. Children were recruited in 5 European countries, ie, Belgium, Denmark, Greece, Sweden, and the United Kingdom. Participation rates ranged from 45% to 96% in the ICSI and IVF groups and from 34% to 78% in the NC group. Cognitive and motor development was assessed with the Wechsler Preschool and Primary Scale of Intelligence-Revised (WPPSI-R) and McCarthy Scales of Children’s Abilities (MSCA) Motor Scale, respectively. The WPPSI-R consists of 2 major scales, ie, Verbal and Performance, each including 6 subtests. The 6 Performance Scale subtests are object assembly, geometric design, block design, mazes, picture completion, and animal pegs. The 6 Verbal Scale subtests are information, comprehension, arithmetic, vocabulary, similarities, and sentences. Scores on the Performance and Verbal Scale subtests are summed to yield the performance IQ (PIQ) and verbal IQ (VIQ), respectively. Scores on both the Performance Scale and the Verbal Scale yield the full-scale IQ (FSIQ). IQ scales have a mean score of 100 and a SD of 15. Each subtest has a mean score of 10 and a SD of 3. The MSCA consists of 6 scales, ie, Verbal, Perceptual-Performance, Quantitative, General Cognitive, Memory, and Motor Scale. In this study, only the Motor Scale was administered. This scale assesses the child’s coordination during performance of a variety of gross- and fine-motor tasks. Leg coordination, arm coordination, and imitative action tests provide measures of gross-motor ability. Draw-a-design and draw-a-child assess fine-motor coordination, as revealed by the levels of hand coordination and finger dexterity. The mean score for this test is 50, with a SD of 10.

Results. No differences were identified among ICSI, IVF, and NC children with respect to VIQ, PIQ, or FSIQ scores of the WPPSI-R. Furthermore, there were no differences between groups regarding the discrepancy between VIQ and PIQ scores. These results were not influenced by gender, country, or maternal educational level. However, in the subgroup of firstborn children with mothers who gave birth at an older age (33–45 years), NC children obtained significantly better VIQ and FSIQ scores than did children conceived through assisted reproductive technologies. These differences in VIQ and FSIQ scores between ICSI/IVF and NC children were relative, because NC children scored <1 IQ point higher than ICSI/IVF children. Therefore, these scores show no clinical relevance. For Verbal Scale subtests, variables such as age of the mother at the time of the birth, educational level of the mother, and gender and nationality of the child interacted with mode of conception, resulting in clinically irrelevant differences between scores for the ICSI/IVF and NC groups on the arithmetic, vocabulary, and comprehension subtests. For Performance Scale subtests, these same demographic factors interacted with mode of conception for the block design, object assembly, and animal pegs subtests, again resulting in clinically irrelevant differences among groups. In the 3 groups (ICSI, IVF, and NC), we observed equal numbers of children scoring below 1 SD from the mean on the WPPSI-R and the MSCA.

Conclusions. This study includes a substantial number of children from several European countries. Apart from a few interaction effects between mode of conception and demographic variables, no differences were found when ICSI, IVF, and NC scores on the WPPSI-R and MSCA Motor Scale were compared. Nevertheless, the aforementioned interaction effects could indicate that demographic variables such as maternal age at the time of the birth and maternal educational level play different roles in the cognitive development of IVF and ICSI chil-
children, compared with NC children. Additional research is needed to explore and verify this finding. Previous studies revealed that ICSI children, in comparison with NC children, more frequently obtained scores below 1 SD from the mean on 3 subtests of the Performance Scale (object assembly, block design, and mazes) or showed a trend of 5.2% of ICSI children, compared with 2.5% of IVF children and 0.9% of NC children, obtaining a score below 1 SD from the mean, but those findings were not confirmed in this study. Here no differences were found among the 3 groups in the numbers of children scoring below 1 SD from the mean on the VIQ, PIQ, and FSIQ tests and the Verbal and Performance Scale subtests. Motor development results were somewhat more conclusive. There were no differences between the scores of ICSI, IVF, and NC children on the MCSA Motor Scale. No interaction effects were found between mode of conception and demographic variables, indicating that these results are not influenced by gender, nationality, maternal educational level, or maternal age at the time of the birth. Furthermore, equal proportions of children in all 3 groups scored below 1 SD from the mean. The results of this study are reassuring for parents who conceived through ICSI (or IVF). The findings indicate that the motor and cognitive development of their offspring is very similar to that of NC children. However, demographic factors such as maternal educational level and maternal age at the time of birth may play different roles in the cognitive development of ICSI and IVF children, compared with NC children. 

Methods

Participants

Children were recruited from 5 different centers in 5 different European countries, ie, Belgium, Denmark, Greece, Sweden, and the United Kingdom. These centers were chosen because of their prominence in the field of assisted reproductive therapies on the European continent. A total of 541 ICSI-conceived children, 440 IVF-conceived children, and 549 NC control children underwent psychologic assessments when they were 4.5 to 5.5 years of age. According to center size, the aim was to recruit equal-sized groups of ICSI, IVF, and NC children, with 175 children per group in Britain and Belgium, 66 children per group in Denmark and Sweden, and 50 children per group in Greece. When targets were not met, additional centers in Flanders and the United Kingdom were contacted. The children were eligible if they were singleton, white, born after ≥32 weeks of gestation, and first or second born, with a native language of English, Dutch, Danish, Swedish, or Greek, respectively. NC control subjects were selected according to the aforementioned criteria and were also matched with respect to maternal education, parental socioeconomic status, gender, and birth order. For the assisted reproductive technology (ART)-conceived group in the United Kingdom, the ICSI children were recruited from a cohort that represented 90% of all children born after ICSI, from 22 of 32 centers that were initially performing ICSI in the United Kingdom. Initial assessment of the children started at 18 months of age. Therefore, these children were recruited from an established cohort. The IVF children were recruited from 5 clinics in the United Kingdom that had established collaborations with the main recruitment center after the initial assessment of ICSI children at 18 months. These IVF children were recruited at age 5 and had not been evaluated previously. The NC children were partly children who had been evaluated in the established cohort and partly newly recruited children from schools in the United Kingdom. The ICSI infants demonstrated a lower mean Bayley Mental Development Index score, compared with both IVF and NC children (96 ± 11, 102 ± 9, and 103 ± 8, respectively). Furthermore, 17% of the ICSI children showed a mild or significant delay in mental development at 1 year of age (Mental Development Index score of <85), compared with only 2% of the children conceived through IVF and 1% of the NC children. The Psychomotor Development Index showed no differences. The same children were evaluated at 5 years of age, by examiners blinded with respect to the mode of conception and with matching for demographic factors. No differences were found in IQ scores between ICSI, IVF, and NC children, and there were no differences in the numbers of children with developmental delays. Unfortunately, psychomotor development was not reassessed in that study. In another controlled study, conducted in Belgium, the Bayley Scales were administered to 201 ICSI children 2 years of age, and no delays in mental development were identified. In the United Kingdom, similar findings were reported when the Griffiths Scales were used to compare 1- to 2-year-old singleton ICSI children with a control group of NC children. Because findings regarding these children’s cognitive and motor development are inconclusive, the aim of this study was to shed more light on the cognitive and motor development of 5-year-old ICSI children.

Findings regarding the cognitive and motor development of ICSI-conceived children are inconclusive. Although it appears there are no major delays in either motor or mental development, larger, concerted, prospective study of many institutions must be undertaken to identify potential minor abnormal-
United Kingdom. In summary, the ICSI-conceived children were evaluated at 18 months; the NC children were mainly seen at 18 months but the IVF children were recruited new at 5 years of age. In Belgium, the majority of the ICSI and IVF children were selected from established cohorts on the basis of consecutive pregnancies in ICSI children and consecutive births, in the case of IVF cycles. The ICSI- and IVF-conceived children participating in this study had been assessed at 2 years of age. To meet the recruitment targets, 5 other centers in Flanders were contacted. ICSI and IVF children recruited from these centers were selected at age 5, like the NC children who were recruited in schools at the same age. In Belgium, 5-year-old children go to nursery schools, which are accessible for all children, ie, normally developing children and mentally delayed children. There are a very few special nursery schools specifically for extremely mentally retarded children or children with specific disabilities, such as visual impairment. ICSI, IVF, and NC children attended mainstream nursery schools. In Sweden, the ART-conceived groups were recruited from complete cohorts of ICSI and IVF children, from registries of deliveries for 2 clinics. The NC control group was selected at birth from the Swedish medical birth registry. Therefore, all children were selected at the time of birth but were not approached until 5 years of age. In Denmark, most ICSI children were recruited from 1 clinic but additional children were recruited from 3 other clinics. Children in all 3 groups (ICSI, IVF, and NC) were recruited from the participating hospital's birth registry. The ICSI and IVF groups consisted of cohorts of children whose mothers had been treated at a fertility clinic. NC control children were selected on the basis of where their mothers gave birth, namely, at the major recruitment center. All children were selected at birth. In Greece, the IVF comparison group was recruited from participating fertility clinics, matched, and assessed in a similar manner as the ICSI-conceived children. Use of the small sample sizes and unequal distribution among groups (ICSI: 26 children; IVF: 13 children; NC: 50 children), data from Greece were not included in the statistical analyses.

Participation Rates

In the United Kingdom, 90% of the ICSI cohort assessed at 18 months of age participated (189 of 201 children).10 Because of ethical restrictions, it was not possible to ascertain response rates for the IVF group and part of the NC group. The NC group had a 50% response rate among the children who were recruited from those seen at 18 months of age. The response rate for the children who were recruited new from schools in the United Kingdom was not known. In Belgium, 45% of eligible children in the ICSI and IVF groups participated and 54% of the children in the NC group who were contacted responded. In Sweden, 96% of ICSI children, 93% of IVF children, and 78% of NC children who were contacted participated. In Denmark, 68% of ICSI children, 56% of IVF children, and 34% of NC children who were contacted participated. In Greece, 25% of ICSI children, 25% of IVF children, and 100% of NC children who were contacted participated. In all countries, the ICSI-conceived children represented consecutive births.

Precautions Against Selection Bias

Because of differing national laws and quality of child health records, strategies for the recruitment of control and comparison groups differed from country to country. In general, strictly applied inclusion criteria, matching for social background, use of standardized test instruments (with an official translated version in every language at each participating center), comparability of training and qualification experience for all examiners, additional training of the examiners (either within the center or together with another team in the United Kingdom) regarding how to assess the children to ensure a uniform protocol, and participation in regular study group meetings during the study period were the precautions taken against selection and interobserver biases. More specifically, there was a very high response rate among the ICSI cohorts in the United Kingdom. This rules out any likely effect from response bias because the rate was 90%. For the IVF and NC groups, there is less confidence regarding response bias but, because of the ethical situation in the United Kingdom, it was not possible to determine details on families who decided not to reply. The schools from which the children were recruited were standard state schools and did not include children who were educationally subnormal. In the United Kingdom, only 1 psychologist, who was blinded with respect to the conception status of all 3 groups, assessed all of the children. Therefore, there was no possibility of intraobserver error being a problem. In Belgium, several psychologists, working closely together, assessed the children. All examiners were evenly involved in the assessments and examined children of all 3 groups, thereby excluding interobserver and intraobserver biases. The psychologists were partially blinded regarding what ART procedure was involved. Because the control group was assessed in schools, blinding for this group was not possible. The ICSI, IVF, and NC children attended mainstream nursery schools and therefore had the same school experience. In Sweden, the 3 groups were selected at birth and all testing was blinded; midwives recruited the patients and sent the names and addresses to the examiners, who were blinded with respect to the mode of conception of the children unless the parents revealed it to them, despite the fact that the parents were told not to give this information. In Denmark, 1 psychologist assessed all children until the last 3 months of the study. During the last 3 months, another examiner, working under her supervision, replaced the original psychologist. Test administration took place at the same location for children of all 3 groups. The 2 psychologists were blinded with respect to the mode of conception. The different recruitment strategies used for each group and country might have introduced some selection bias. The ideal situation would have been to have a national birth registry in each country. In the absence of a national birth registry, the child's socioemotional development.14 In addition to a psychologic examination, a pediatric examination was conducted.

Procedure

Children's cognitive and motor development was assessed with the Wechsler Preschool and Primary Scale of Intelligence-Revised (WPPSI-R)11 and the McCarthy Scales of Children's Abilities (MSCA) (Motor Scale),12 respectively. Two additional items were administered to gather more information on hand preference. All children were individually assessed by a psychologist.

Children were also presented with a procedure designed to evaluate their perceptions of their family members with the Family Relation Test.13 Parents were asked to fill out a number of questionnaires measuring parental mental health, family functioning, and the child's socioemotional development.14 In addition to a psychologic examination, a pediatric examination was conducted.

Outcome Measures

The WPPSI-R

The WPPSI-R is an individually administered, clinical instrument for assessing the intelligence of children of age 3 years to 7 years 3 months. The WPPSI-R provides standardized measures of a variety of abilities reflecting different aspects of intelligence. It consists of 2 major scales, ie, Verbal and Performance, each including 6 subtests. The 6 Performance Scale subtests are object assembly, geometric design, block design, mazes, picture completion, and animal pegs. The 6 Verbal Scale subtests are information, comprehension, arithmetic, vocabulary, similarities, and sentences. Scores on the Performance and Verbal Scale subtests are summed to yield the performance IQ (PIQ) and the verbal IQ (VIQ), respectively. Scores on both the Performance Scale and the Verbal Scale yield the full-scale IQ (FSIQ).11 IQ scales have a mean score of 100 and a SD of 15. Each subtest has a mean score of 10 and a SD of 3.

The MSCA Motor Scale

The MSCA is appropriate for children from 2.5 through 8.5 years of age. The MSCA consists of 6 scales, ie, Verbal, Perceptual-Performance, Quantitative, General Cognitive, Memory, and Motor. In this study, only the Motor Scale was administered. This scale assesses the child's coordination during performance of a variety of gross- and fine-motor tasks. Leg coordination, arm coordination, and imitative action tests provide measures of gross-motor ability. Draw-a-design and draw-a-child assess fine-motor coordination, as revealed by the levels of hand coordination and finger dexterity. The mean score for this test is 50, with a SD of 10.
Observations of Hand and Eye Preference With the MSCA Motor Scale

Several items of the Motor Scale were also used to assess selected aspects of laterality. Four observations of hand dominance are made during the administration of the Motor Scale. The examiner notes the hand the child uses for bouncing a ball, catching a beanbag, throwing a beanbag, and drawing. The child then may be classified as “hand dominance established” (right- or left-handed) or “hand dominance not established.” The child’s eye preference is noted as sight through a tube. In addition, 2 items, namely, “comb your hair” and “pretend to eat with this spoon,” were administered. The examiner noted which hand was used to perform both of these actions.

Statistical Analyses

The VPSSI-R and the MSCA Motor Scale are standardized tests with established psychometric (reliability and validity) properties.11,12 FIQ, PIQ, and VIQ scores obtained with the VPSSI-R and MSCA Motor Scale scores were considered as dependent variables with an interval-measurement scale. With SPSS software (version 11; SPSS, Chicago, IL), 3-way analyses of variance were conducted with interval data obtained with the VPSSI-R and the MCSA Motor Scale and the following independent variables: mode of conception, country, gender, educational level of the mother, and age of the mother at the time of the birth. Dichotomous variables such as hand and eye dominance and use of the spoon and comb were analyzed with χ² tests. A significance level of 5% (2-tailed) was accepted throughout.

Role of the Funding Source

The protocol was approved by the ethics committee of each institution, in accordance with national regulations in each country. The European Union 5th Framework Quality of Life Program financed this project; the funding source had no responsibility for the study design or the interpretation of data.

RESULTS

Demographic Features

Data for a total of 511 ICSI-conceived children, 424 IVF-conceived children, and 488 NC children were analyzed. The average ages at testing were similar for the groups (ICSI: 5.05 years; IVF: 5.02 years; NC: 5.06 years; range: 4.20–6.40 years). All 3 groups included equal numbers of boys and girls, but there were significantly fewer firstborn children in the NC control group (χ² = 28.01, P < .001).

There were significant differences between the conception groups with respect to maternal and paternal age. Parents of ICSI and IVF children were significantly older than parents of NC children but were similar to each other [maternal age: ICSI: 33.5 years; IVF: 33.8 years; NC: 31.3 years; F(2,1334) = 49.99, P < .001; paternal age: ICSI: 35.5 years; IVF: 35.5 years; NC: 32.7 years; F(2,1249) = 44.23, P < .001]. Educational levels of mothers and fathers were defined with the Hollingshead Social Class Index.15 Then parental education was categorized into 3 levels, high (college graduate), medium (attended college or is a high school graduate), and low (passed only part of high school or less). Mothers (n = 318) and fathers (n = 379) with high educational levels had obtained a postgraduate degree or a graduate degree. Mothers (n = 687) and fathers (n = 577) with medium educational levels had entered a university or had passed school matriculation fully. Mothers (n = 315) and fathers (n = 360) with low educational levels had passed school matriculation partially or had no qualification at all.

More mothers in the control group had high levels of education and fewer had low educational levels (χ² = 28.88, P < .001). No significant differences were found in educational levels among the fathers in the ICSI, IVF, and NC groups. Maternal occupation and paternal occupation, defined with the Classification of Occupations,16 were not associated with conception group.

Cognitive Development

No differences were identified among ICSI, IVF, and NC children with respect to VPSSI-R VIQ, PIQ, and FSQ scores (Table 1). Furthermore, there were no differences among groups regarding the discrepancy between VIQ and PIQ scores. These results were not influenced by gender, country, or educational level. However, interaction effects were observed for age of the mother at the time of the birth and VIQ and PIQ scores. Firstborn children in the ICSI and IVF groups with mothers between 33 and 45 years of age demonstrated lower VIQ and PIQ scores, compared with their NC counterparts [F(2,1101) = 3.36, P < .05, and F(2,1101) = 3.53, P < .05, respectively].

Verbal Scale Subtests

A 3-way analysis of variance with mode of conception, gender, and country as independent variables revealed a conception mode-country effect on the Verbal Scale subtest arithmetic. In Denmark, ICSI children scored significantly lower than children in the IVF and NC groups [F(6,1355) = 2.13, P = .048]. ICSI and IVF boys obtained lower scores on the vocabulary subtest, in comparison with NC boys and ICSI-conceived, IVF-conceived, and NC girls [F(2,1360) = 3.53, P = .029].

An interaction effect between mode of conception and educational level of the mother was found, showing that IVF children of mothers with high educational levels scored significantly higher on the
vocabulary subtest than did ICSI and NC children of mothers with similar educational levels [$F(4,1095) = 3.76, P < .05$]. For the comprehension subtest, the age of the mother at the time of the birth influenced children's scores. In the NC group, scores for the comprehension subtest increased when the mother had her child at an older age (33–45 years); in the ICSI and IVF groups, the opposite was true [$F(2,1092) = 5.96, P < .05$] (Table 2).

**Performance Scale Subtests**

A significant conception mode-country effect was found, indicating that in Belgium NC children scored higher on the block design subtest, whereas in Denmark the opposite was true [$F(6,1355) = 4.00, P = .001$]. Also, for the block design subtest, NC children with low-educated mothers scored higher than ICSI and IVF children with equally low-educated mothers, whereas ICSI children with mothers with medium levels of education scored better than IVF and NC children with medium-educated mothers [$F(4,1096) = 3.33, P < .05$]. For the object assembly subtest, the educational level of the mother and the conception mode interacted, with NC children of highly educated mothers scoring lower on this subtest, in comparison with ICSI and IVF children [$F(4,1098) = 2.39, P < .05$]. For this subtest, another interaction effect was revealed between conception mode and age of the mother at the time of the birth. In the NC group, the scores of children of mothers who gave birth at older ages (33–45 years) were significantly higher than the scores of children of young mothers (19–32 years), in contrast to the scores of children in the IVF group [$F(2,1100) = 3.00, P = .05$]. In the animal pegs subtest, a conception mode-gender-country interaction was revealed. NC boys in Denmark obtained higher scores than did ICSI boys, whereas Belgian IVF girls scored better than ICSI girls (Table 3).

**Marked Developmental Delay**

There were no conception group differences in the extent to which children showed marked developmental delay, defined as scoring below 1 SD from the mean on VIQ, PIQ, or FSIQ tests or Verbal and Performance Scale subtests (Table 4). Discrepancies of ≥12 points between VIQ and PIQ scores might be indicative of learning disabilities. No difference was found in the number of children in the ICSI, IVF, and NC groups with a discrepancy of ≥12 points between VIQ and PIQ scores [$\chi^2(4) = 3.08$, not significant].

**Motor Development**

No significant differences were found between ICSI, IVF, and NC children’s average scores on the MCSA Motor Scale. Motor development was not influenced by the educational level of the mother, the age of the mother at the time of the birth, gender, or country. Also, no differences were found in the proportions of children in the 3 conception groups scoring below 1 SD from the mean on the Motor Scale (score of <40) (Table 4). When hand dominance and eye dominance were examined, there were no differences among groups [$\chi^2(6) = 6.67$, not significant, and $\chi^2(4) = 5.97$, not significant, respectively]. This finding was confirmed by the results of complementary assessments in which the children had to demonstrate how they handled a spoon and a comb [$\chi^2(4) = 1.68$, not significant, and $\chi^2(4) = 1.77$, not significant, respectively].

**DISCUSSION**

This study includes a substantial number of 5-year-old children assessed for cognitive and motor development in 4 European countries. No differences were identified between ICSI, IVF, and NC children on the VIQ, PIQ, or FSIQ scores of the WPPSI-R. Furthermore, there were no differences between groups regarding the discrepancy between VIQ and PIQ scores. These results were not influenced by gender, country, or maternal educational level. However, in the subgroup of firstborn children with mothers who gave birth at an older age (33–45 years), NC children obtained significantly better VIQ and FSIQ scores than did children born after ART. These differences in VIQ and FSIQ scores between ICSI/IVF children and NC children are relative, because NC children scored <1 IQ point higher than the ICSI/IVF children. Therefore, these scores show no clinical relevance.

For the Verbal Scale subtests, variables such as age of the mother at the time of the birth, educational level of the mother, and gender and nationality of the children interacted with the mode of conception, resulting in clinically irrelevant differences between

---

**TABLE 2.** Mean Verbal Scale Test Scores, Measured With the WPPSI-R, According to Conception Mode

<table>
<thead>
<tr>
<th>Score</th>
<th>ICSI Mean ± SD</th>
<th>IVF Mean ± SD</th>
<th>NC Mean ± SD</th>
<th>Statistical Analysis*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information (n = 1363)</td>
<td>11.70 ± 2.78</td>
<td>11.37 ± 2.29</td>
<td>11.67 ± 2.73</td>
<td>NS</td>
</tr>
<tr>
<td>Comprehension (n = 1358)</td>
<td>10.86 ± 2.69</td>
<td>10.88 ± 2.39</td>
<td>10.95 ± 2.51</td>
<td>P &lt; .05‡</td>
</tr>
<tr>
<td>Arithmetic (n = 1363)</td>
<td>10.24 ± 2.47</td>
<td>10.62 ± 2.60</td>
<td>10.54 ± 2.54</td>
<td>P &lt; .05‡</td>
</tr>
<tr>
<td>Vocabulary (n = 1363)</td>
<td>12.42 ± 2.79</td>
<td>12.39 ± 2.79</td>
<td>12.41 ± 2.75</td>
<td>P &lt; .05§</td>
</tr>
<tr>
<td>Similarities (n = 1360)</td>
<td>11.20 ± 3.50</td>
<td>10.99 ± 3.27</td>
<td>11.57 ± 3.28</td>
<td>NS</td>
</tr>
</tbody>
</table>

* Three-way analysis of variance (conception mode × country × gender and conception mode × educational level × age of the mother at the time of the birth). NS indicates nonsignificant ($P > .05$).
‡ Significant conception mode-age of the mother at the time of the birth interaction effect.
§ Significant conception mode-country interaction effect.
§§ Significant conception mode-gender interaction effect and conception mode-educational level of the mother interaction effect.
TABLE 3. Mean Performance Scale Subtest Scores, Measured With the WPPSI-R, According to Conception Mode

<table>
<thead>
<tr>
<th>Subtest</th>
<th>ICSI</th>
<th>IVF</th>
<th>NC</th>
<th>Statistical Analysis*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td></td>
</tr>
<tr>
<td>Object assembly (n = 1366)</td>
<td>11.30 (2.90)</td>
<td>11.30 (3.03)</td>
<td>11.17 (2.87)</td>
<td>P &lt; .05†</td>
</tr>
<tr>
<td>Block design (n = 1364)</td>
<td>11.16 (3.33)</td>
<td>10.77 (3.10)</td>
<td>10.94 (2.99)</td>
<td>P &lt; .05‡</td>
</tr>
<tr>
<td>Mazes (n = 1365)</td>
<td>10.09 (3.46)</td>
<td>9.81 (3.47)</td>
<td>9.58 (3.39)</td>
<td>NS</td>
</tr>
<tr>
<td>Picture completion (n = 1366)</td>
<td>11.09 (2.57)</td>
<td>10.62 (2.65)</td>
<td>11.10 (2.66)</td>
<td>NS</td>
</tr>
<tr>
<td>Animal pegs (n = 1355)</td>
<td>11.00 (2.96)</td>
<td>11.21 (2.61)</td>
<td>11.14 (2.67)</td>
<td>P &lt; .05§</td>
</tr>
</tbody>
</table>

* Three-way analysis of variance (conception mode x country x gender and conception mode x educational level x age of the mother at the time of the birth). NS indicates nonsignificant (P > .05).
† Significant conception mode-country-gender interaction effect.
‡ Significant conception mode-country interaction effect and conception mode-educational level of the mother interaction effect.
§ Significant conception mode-country-gender interaction effect.

TABLE 4. Proportions of Children Scoring Below 1 SD From the Mean on VIQ, PIQ, and FSIQ Tests of the WPPSI-R (<85) and the WPPSI-R Subtests (7) and of Children Scoring Below 1 SD From the Mean on the MCSA Motor Scale (<40)

<table>
<thead>
<tr>
<th>Subtest</th>
<th>No. (%)</th>
<th>Statistical Analysis*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ICSI</td>
<td>IVF</td>
</tr>
<tr>
<td>VIQ</td>
<td>14/483 (2.9)</td>
<td>7/416 (1.7)</td>
</tr>
<tr>
<td>PIQ</td>
<td>31/483 (6.4)</td>
<td>25/415 (6.0)</td>
</tr>
<tr>
<td>FSIQ</td>
<td>21/483 (4.3)</td>
<td>17/399 (4.3)</td>
</tr>
<tr>
<td>Comprehension</td>
<td>12/483 (2.5)</td>
<td>5/413 (1.2)</td>
</tr>
<tr>
<td>Arithmetic</td>
<td>14/482 (2.9)</td>
<td>10/412 (2.4)</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>11/483 (2.3)</td>
<td>10/413 (2.4)</td>
</tr>
<tr>
<td>Similarities</td>
<td>48/481 (10)</td>
<td>35/412 (8.5)</td>
</tr>
<tr>
<td>Object assembly</td>
<td>24/485 (4.9)</td>
<td>22/414 (5.3)</td>
</tr>
<tr>
<td>Block design</td>
<td>32/484 (6.6)</td>
<td>41/413 (9.9)</td>
</tr>
<tr>
<td>Mazes</td>
<td>58/485 (1.2)</td>
<td>69/414 (1.7)</td>
</tr>
<tr>
<td>Picture completion</td>
<td>22/485 (4.5)</td>
<td>23/414 (5.6)</td>
</tr>
<tr>
<td>Animal pegs</td>
<td>45/481 (9.4)</td>
<td>27/407 (6.6)</td>
</tr>
<tr>
<td>MSCA Motor Scale Index</td>
<td>47/479 (9.8)</td>
<td>32/409 (7.8)</td>
</tr>
</tbody>
</table>

* Pearson χ² test. NS indicates nonsignificant (P > .05).

scores of the ICSI/IVF and NC groups on the arithmetic, vocabulary, and comprehension subtests. For the Performance Scale subtests, the same demographic factors interacted with the mode of conception for the block design, object assembly, and animal pegs subtests, again resulting in clinically irrelevant differences between groups. However, the aforementioned interaction effects could indicate that demographic variables such as maternal age at the time of the birth and maternal educational level play different roles in cognitive development of IVF and ICSI children, compared with NC children. Additional research is needed to explore and verify this finding.

In general, the few controlled studies of mental and/or motor development of ICSI-conceived children at the age of 5 provided similar results.6,7 One study7 revealed no significant differences in WPPSI-R VIQ, PIQ, and FSIQ scores. Moreover, the mean results for all 3 IQ scales were in the normal range for all 3 groups, and no between-group differences were found for any of the scales. As in previous studies of the psychologic development of 5-year-old ICSI children,7,17 the number of children in each group who scored below 1 SD from the mean on the main scale and subscales (Verbal and Performance) of the WPPSI-R were compared. Those studies indicated that ICSI children, in comparison with NC children, more frequently obtained scores below 1 SD from the mean on 3 subtests of the Performance Scale (object assembly, block design, and mazes subtests)17 or showed a trend (P = .19) of 5.2% of ICSI children, compared with 2.5% of IVF children and 0.9% of NC children, obtaining scores below 1 SD from the mean,7 but those findings were not confirmed in this study. Here no differences were found among the 3 groups in the numbers of children who scored below 1 SD from the mean on the VIQ, PIQ, and FSIQ tests and Verbal and Performance Scale subtests. Results regarding motor development were somewhat more conclusive. There were no differences in the MCSA Motor Scale scores for ICSI, IVF, and NC children. No interaction effects were found between the mode of conception and demographic variables, indicating that these results were not influenced by gender, nationality, maternal educational level, or maternal age at the time of the birth. Furthermore, there were equal proportions of children in all 3 groups who scored below 1 SD from the mean.

The results of this study are reassuring for parents who have conceived through ICSI (or IVF). The findings indicate that the motor and cognitive development of their offspring is very similar to that of NC children. However, demographic factors such as maternal educational level and maternal age at the time of the birth might play different roles in the cognitive development of IVF and ICSI children, compared with NC children. Additional research is needed to explore and verify this finding.
development of ICSI and IVF children, compared with NC children.

ACKNOWLEDGMENTS

The European Union 5th Framework Quality of Life Program (contractQLG4-CT-2000-00545) financed this project (International Collaborative Study of ICSI: Child and Family Outcomes).

We thank the families involved and all of the fertility clinics and schools that helped us with subject recruitment. The fertility centers of University Clinic Ghent, Academic Medical Clinic Middelheim Antwerp, Academic Medical Clinic St Jan Bruges, and Fertility Clinic Genk are kindly acknowledged for their contribution of some of the ICSI and IVF patients in the Belgian group. The fertility centers of CARE and NURTURE (Nottingham), Kings College (London), London Fertility Centre, Leeds IVF Unit, Chelsea and Westminster IVF Unit (London), and University of Hull are kindly acknowledged for contribution of some of the ICSI and IVF patients in the United Kingdom group. Other centers assisted on previous occasions with recruitment of the original United Kingdom cohort and provision of clinic space. The following fertility centers of Athens and Thessaloniki are kindly acknowledged for contribution of some of the ICSI and IVF patients in the Greek group: Center for Human Reproduction (Athens), AKESO Center (Athens), Iatriki Erevna (Athens), Fertility Institute (Athens), and Embryogenesis Center (Athens). The midwives Kia Borg and Gunilla Borg, who assisted with recruitment in Sweden, are also thanked. The Fertility Clinic at Herlev University Hospital, the Fertility Clinic Trianglen (Copenhagen), and the Danish Fertility Clinic (Copenhagen) are acknowledged for assistance in recruiting ICSI children in the Danish group.

REFERENCES

7. Leslie GL, Gibson FL, McMahon C, Cohens J, Saunders DM, Tennant C. Children conceived using ICSI do not have an increased risk of delayed mental development at 5 years of age. Hum Reprod. 2003;18:2067–2072
15. Hollingshead A. Four-Factor Index of Social Status. New Haven, CT: Yale University; 1975
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
DOI: 10.1542/peds.2004-1445

The online version of this article, along with updated information and services, is located on the World Wide Web at:
/content/115/3/e283.full.html