Journal Pre-proof

Assisted reproductive technology and neurodevelopmental outcomes in offspring: a prospective birth cohort study in East China

Wei Li , Jiuru Zhao , Meng Ni , Qianqian Zhang , Qianwen Shen , Hong Li , Zheng Tang , Dongting Yao , Tao Wang , Sudong Qi , Baihe Li , Xiya Ding , Jinliang Xie , Xiaojin Wang , Bingshun Wang , Zhiwei Liu

PII: S1472-6483(23)00112-8
DOI: https://doi.org/10.1016/j.rbmo.2023.02.006
Reference: RBMO 3180

To appear in: *Reproductive BioMedicine Online*

Received date: 7 November 2022
Revised date: 19 January 2023
Accepted date: 13 February 2023

Please cite this article as: Wei Li , Jiuru Zhao , Meng Ni , Qianqian Zhang , Qianwen Shen , Hong Li , Zheng Tang , Dongting Yao , Tao Wang , Sudong Qi , Baihe Li , Xiya Ding , Jinliang Xie , Xiaojin Wang , Bingshun Wang , Zhiwei Liu , Assisted reproductive technology and neurodevelopmental outcomes in offspring: a prospective birth cohort study in East China, *Reproductive BioMedicine Online* (2023), doi: https://doi.org/10.1016/j.rbmo.2023.02.006

Assisted reproductive technology and neurodevelopmental outcomes in offspring: a prospective birth cohort study in East China

Wei Li[1,2,3], Jiuru Zhao[1,2,3], Meng Ni [1,2,3], Qianqian Zhang[1,2,3], Qianwen Shen [1,2,3], Hong Li[2], Zheng Tang[2], Dongting Yao[1,2,3], Tao Wang[1,2,3], Sudong Qi[1,2,3], Baihe Li[1,2,3], Xiya Ding[1,2,3], Jinliang Xie[4], Xiaojin Wang[4], Bingshun Wang[4], Zhiwei Liu[1,2,3]*

[1]International Peace Maternity and Child Health Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China

[2]International Peace Maternity and Child Health Hospital of China Welfare Institution, School of Medicine, Shanghai Jiao Tong University, Shanghai, China

[3]Shanghai Key Laboratory of Embryo Original Diseases, Shanghai, China

[4]Department of Biostatistics, Clinical Research Institute, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China.

***Corresponding author:**

Zhiwei Liu

Department of Neonatology, International Peace Maternity and Child Hospital,

School of Medicine, Shanghai Jiao Tong University, 910# Hengshan Road, Shanghai 20030, China.

Fax: 86-21-64073899

Tel: 86-21-64073899

E-mail: liuzhiwei@hotmail.com

ABSTRACT

Research question: Is ART associated with adverse neurodevelopmental outcome in 12-month-old offspring compared with those conceived through spontaneous conception (SC)?

Design: In this prospective cohort study, 488 infertile women undergoing ART and 1397 women with SC were recruited and followed until their offspring were 12 months old. The primary outcome was neurodevelopment in the offspring. The association between exposure to ART and Gesell developmental scale scores was investigated using multiple linear regression models after

adjusting for confounders. Propensity score matching (PSM) and inverse probability of treatment weight (IPTW) were used to verify the results.

Results: In total, 18 (3.7%) and 40 (2.9%) children in the ART and SC groups, respectively, had been diagnosed with neurodevelopmental delay at 12 months of age. We found that gross motor, adaptive behavior, language, and total development quotient (DQ) scores were to be comparable between the groups. Following multivariate linear regression and IPTW, social behavior DQ scores were found to be slightly higher in the ART group than in the SC group. Higher social behavior DQ scores in the ART group were also observed in males and in singleton subgroup.

Conclusions: At 12 months, offspring born after ART appeared to have similar motor, language, and adaptive behavior skills, and total DQ scores to those born after SC. However, social behavior development in 12-month-old infants was slightly higher in those conceived using ART than in those with SC, especially in male and in singleton infants. Our findings may provide new information in evaluating the potential benefits and risks of ART treatment.

Keywords: Assisted reproductive technology, infertility characteristic, offspring neurodevelopmental outcome, social behaviour skills

Key Message

1. Offspring born after ART appeared to have similar motor, language, and adaptive behavior skills, and total DQ scores to those born after SC at one year old.

2. Social behavior development in 12-month-old infants was slightly higher in those conceived using ART than in those with SC, especially in male and in singleton infants.

1. INTRODUCTION

In China, assisted reproductive technology (ART), mainly involving in vitro fertilisation (IVF) and intracytoplasmic sperm injection (ICSI), has been widely used for the treatment of infertility since first introduced in 1988 (Zhang et al. 2021). Globally, it is estimated that >8 million infants have been born with the use of ART (Barbuscia et al. 2020), with >1% of those infants born in mainland China (Yang et al. 2014). Given the increasing number of children conceived using ART,

several studies have investigated the effect of ART on offspring wellbeing throughout childhood. Some international and Chinese studies have reported that children born after ART are at an elevated risk of adverse health outcomes, such as birth defects (Lv et al. 2021, Boulet et al. 2016, Wang et al. 2021), metabolic diseases (Cui et al. 2020, Guo et al. 2017), and cancer (Hargreave et al. 2019) compared with those born after spontaneous conception (SC). However, studies that have investigated the association between exposure to ART and neurodevelopmental outcomes in children have reported inconsistent findings. Several studies have reported an association between the use of ART and an increased risk of intellectual disability (Hansen et al. 2018) and autism spectrum disorders (Sandin et al. 2013, Halpert 2012), whereas various other studies and meta-analyses have found that children conceived using ART had similar cognitive, motor, and language development as children born after SC (Balayla et al. 2017, Bay, Mortensen and Kesmodel 2013). In China, there are limited data concerning neurodevelopmental outcomes in children born after ART.

Studies that have reported outcomes for children born after ART suggest that multiple factors may influence child growth and development (Bay et al. 2013, Roychoudhury et al. 2021), including parental demographics such as age, complications during pregnancy (Zhu et al. 2009, Roychoudhury et al. 2021), underlying subfertility or infertility factors, and aspects of the ART treatment itself (Goisis et al. 2019, Wang et al. 2017, Pontesilli et al. 2021). ART procedures may also increase the risk of neurodevelopmental abnormalities through promoting elevated risks of pregnancy-related complications and adverse birth outcomes, which in turn lead to a higher risk of poor neurologic outcomes. Specifically, preterm birth and low birth weight (LBW) are important risk factors for neurodevelopmental impairment in infants (Blencowe et al. 2013, Scharf et al. 2016). Moreover, ART can lead to a higher rate of multiple births and is a major contributor to the elevated prevalence of preterm birth and LBW in both single and twin pregnancies (Lin et al. 2021, Qin et al. 2016, Declercq et al. 2015). Therefore, it is necessary to evaluate these associations with consideration to several potential factors. One study indicated that offspring development in early childhood differed between male and female children (Donald et al. 2019). The influence of sex differences on neurological development among ART offspring remains unclear and requires further investigation.

Given the large number of factors that may influence the neurodevelopmental performance of ART offspring, robust research methods are required to examine these associations. ART-related specific adverse birth outcomes and parent-related factors should be carefully evaluated and adjusted for when analysing offspring outcomes. We undertook a longitudinal study in Shanghai, China, to compare women who received ART with those who conceived spontaneously. We followed up their offspring until they reached 12 months of age. We aimed to investigate the association between ART exposure and neurodevelopmental outcomes in 12-month-old offspring compared with those conceived through SC.

2. PATIENTS and METHODS

2.1 Study design and population

This prospective cohort study was conducted as a sub-cohort of the China National Birth Cohort study (CNBC), which is recruiting in Shanghai, East China. This cohort is initially based on a 5-year recruitment (2016-2021) and is still following-up assessment. The study recruited women who were about to receive ART treatment and women who were in the first trimester of SC (controls) at the Centre for Reproductive Medicine at the International Peace Maternity and Child Health Hospital (IPMCH) from July 1, 2016, to December 31, 2019. About half of participants had delivered between this period, and offspring follow-up was completed a year later. The inclusion criteria comprised women: (i) who were residents of Shanghai; (ii) who were able to read and write in Chinese; (iii) who had no severe history of mental illness; and (iv) who received antenatal care and delivery at IPMCH. For the ART group, couples eligible for ART were those aged at least 20 years, who had not become pregnant after a year or more of unprotected intercourse, and who had been seeking fertility evaluation and treatment. In this study, ART mainly refers to traditional IVF and ICSI. The control group comprised: pregnant women between 8 and 14 weeks of gestation who were planning a delivery at IPMCH and who volunteered to participate in this study. Local reproductive professionals or obstetricians screened parents for eligibility and explained the study design, including the purpose and procedures of the study, the potential harms and benefits, and confidentiality issues. All children were followed up regularly at 42 days, and at 6 and 12 months postpartum. The last follow-up date was August 30, 2021. Briefly, mothers were interviewed via telephone or WeChat using questionnaires related to the

health of both the mothers and their children at 42 days and at 6 months postpartum. Once cohort offspring had reached approximately 12 months of age, or were within 60 days of the target age, they received an invitation to attend the IPMCH for a systematic medical examination. The follow-up rate for infants aged 12 months was >90%.

Neonates with major congenital or chromosomal anomalies, women who had miscarriages and stillbirths, and those lost to follow-up were excluded. Women who used donor oocytes were also ineligible for enrolment due to ethical requirements of anonymity and confidentiality. Finally, complete data on the assessment of offspring neurodevelopment was available for 1397 children conceived after ART and for 488 children born after SC. A flow chart of the study participants is shown in **Figure 1**. Five hundred and thirteen offspring (27%) of SC and one hundred and twenty-two offspring (20%) of ART group were excluded. To determine the potential for bias, we compared the baseline characteristics between the follow-up population and those who were excluded. However, the baseline characteristics between those followed and those excluded were not significantly different.

2.2 Data collection and ethical approval

Demographic characteristics and clinical data were prospectively collected using face-to-face questionnaires and electronic medical records.

All methods and study protocols were approved by the Medical Ethics Committee of the International Peace Maternity and Child Health Hospital, affiliated with Shanghai Jiao Tong University School of Medicine (approval number: GKLW-2016-21). Each participant provided informed consent in accordance with the Declaration of Helsinki.

2.3 Exposure

ART was the primary exposure in this study. Participants in the ART group received either IVF or ICSI therapy. Parental demographic characteristics such as age, education level, household income, pre-pregnancy body mass index (BMI), diseases during pregnancy, tobacco smoking (yes/no), alcohol consumption (yes/no), plurality, infants' sex, gestational age, and birth weight affect the risk of neurodevelopmental outcomes in offspring, which we then considered when exploring the association between exposure factors and outcome variables.

2.4 Neurodevelopmental assessment

Trained healthcare professionals obtained authorisation to use the Gesell developmental scale (GDS, Chinese version) assessment and were blinded to participants' grouping and baseline information. Neurodevelopmental assessments were undertaken at a specialised clinic and children were accompanied by their parents. The GDS has been widely used to evaluate the neurodevelopment of children aged 0-3 years and has been found to have adequate concurrent and discriminant validity, and high reliability (Dror et al. 2009, Wang et al. 2018). The Chinese version of the GDS, which was adapted to the Chinese population by the Chinese Pediatric Association, was administered to infants at 12 months of age in this study. The GDS divided the children's neurodevelopment into five areas: gross motor skills, fine motor skills, language, adaptive behaviour, and social behaviour. Each child was assigned a development quotient (DQ) in each of the five domains. The total DQ of the five domains was used to evaluate the overall level of neurodevelopment, with higher DQs indicating greater neurodevelopment. Children with normal development had a total DQ score >85 on the GDS. Neurodevelopmental delay (ND) was defined as a DQ score below the normal threshold.

2.5 Statistical methods

The main analysis was undertaken in four steps. First, we used descriptive statistics to summarise the parental demographic characteristics and the clinical birth outcomes of the offspring. Subsequently, we compared offspring neurodevelopmental outcomes in the ART group with those in the SC group. Second, we applied multiple linear regression models to explore the association between ART and offspring neurodevelopment with correction for confounders. Third, since baseline clinical characteristics were considerably unbalanced between the two groups, two different matching methods were used to reduce the effect of confounders when analysing the data. Finally, to ensure robustness of the primary study, we conducted several other sensitivity analyses.

A Kolmogorov-Smirnov test confirmed that the continuous variables were normally distributed. The continuous variable data were expressed as mean ± standard deviation (SD) unless otherwise specified, according to the distribution. Categorical variables were expressed as numbers (n) with percentages (%). Depending on whether data were normally distributed, the Student's t-test or the Mann–Whitney U test was used to compare continuous variables. Pearson's chi-square test or

Fisher's exact test (two-tailed) was used for categorical variables in the two groups.

Multiple linear regression models were utilized to evaluate the associations between ART exposure and infant outcomes in terms of neurodevelopment, which were adjusted for parental age, maternal educational levels, maternal pre-pregnancy BMI, household income, alcohol and tobacco use during pregnancy, plurality, and infants' sex in adjusted Model, except the variable used for stratified. Potential confounding variables were selected from previous studies of ART exposure on offspring neurological outcomes, as well as being based on the statistical considerations of the present study.

For the primary outcome, two models of the association between DQ and treatment modality were then constructed, namely, a propensity score model (PSM) and an inverse probability of treatment weight (IPTW) model. The propensity score of receipt was constructed via a multivariable logistic regression model, which included variables significantly associated with treatment modality via univariable analysis and variables with significant importance clinically. Based on the propensity score, an IPTW was calculated and truncated at the 1st and 99th percentiles. To construct the PSM, those in the ART group were matched 1:2 with those in the SC group on the propensity score using the greedy, nearest□neighbor matching algorithm, and a maximum allowed difference of ±10% for propensity scores was used. Furthermore, uni- and multivariable analyses were performed to assess the effects of ART on the children's DQ.

For subgroup analysis, we compared the five domains and total DQ scores between the ART and SC groups and stratified them according to plurality (singleton and multiplets) or infants' sex (male and female), after adjusting for confounders. We also considered that birth outcomes (preterm and LBW) may have had an adverse effect on neurodevelopment. The sensitivity analyses concerning the infants' neurodevelopmental outcomes in relation to full-term and normal birth weight births are also calculated in the current study**.**

All statistical analyses were performed using SPSS 26.0 (IBM SPSS Statistics, IBM Corporation, Armonk, NY) and R statistical (version 4.0.6, R Foundation for Statistical Computing, Vienna, Austria) software. Statistical significance was set at a two-tailed *P*-value <0.05.

3. RESULTS

3.1 Baseline characteristics

Our final cohort comprised 1885 parent-offspring pairs. We compared 1397 pregnancies after SC with 488 pregnancies after ART. The baseline characteristics of all the parents and offspring are listed in **Table 1**. Regarding parents' sociodemographic characteristics, women who underwent ART were more likely to be >35 years old, multiparous, with a higher pre-pregnancy BMI, a lower education level, and having engaged in tobacco use during pregnancy compared with those aged <35 years, but less likely to have a higher household income. Fathers who underwent ART were more likely to be >35 years old and have a lower education level. Additionally, the prevalence of diabetes (chronic or gestational diabetes) and preeclampsia was considerably higher among offspring in the ART group. In terms of birth outcomes, the ART group was more likely than the SC group to have higher proportions of multiple births (25.6% vs. 1.0%), LBW (birthweight <2,500 g; 14.0% vs. 2.9%) and preterm births (gestational age, <37 weeks; 20.3% vs. 5.4%), and were more likely to be born by caesarean delivery (73.0% vs. 41.2%), respectively.

3.2 Primary outcome

In the study cohort, mean (SD) scores for offspring total DQ were 96.81 (SD, 6.24) in the SC group and 96.55 (SD, 6.25) in the ART group. Furthermore, 3.7% (n = 18) of the ART group and 2.9% (n = 40) of the SC group had been diagnosed with ND. The primary outcomes are shown in **Table 2**. Nonparametric Kruskal-Wallis test and univariate linear regression analysis (crude model) results indicated that there were no significant differences in developmental scores in gross motor and fine motor skills, language, and adaptive and social behaviour domains between the two groups. However, after multivariate linear regression analysis, ART was found to be significantly positively associated with social behaviour DQ scores (Adjusted Model, adjusted b = 1.08; 95% confidence interval [CI] 0.10–2.06; P = 0.031) (**Table 3**), after adjusting for paternal age, maternal age, maternal education, income, maternal pre-pregnancy BMI, smoking, alcohol consumption during pregnancy, plurality, and infants' sex.

To mitigate potential bias due to discrepancies in baseline covariates between the two groups, we applied two different matching methods, namely, IPTW and PSM. The imbalance in covariates between the two groups was significantly reduced after applying these methods. The baseline characteristics of the study populations (parents and offspring) after applying IPTW and PSM are

presented in **Table 1 and Supplementary Table S1,** respectively. A higher social behaviour score in the ART group was still evident after using the multiple linear regression model in IPTW-adjusted analysis (adjusted b = 0.99; 95% CI 0.26–1.71; P = 0.008) (**Table 3**). In both the IPTW and PSM models, no statistical difference in gross motor skills, adaptive behaviour, language, and total DQ scores was found between the two groups. Nevertheless, the IPTW and PSM cohort incidentally showed that the mean (SD) score concerning offspring fine motor DQ in the ART group was lower than that in the SC group (IPTW: 96.94 [6.61] vs. 96.33 [6.22], P = 0.002; PSM: 96.97 [6.58] vs. 96.56 [6.59], P < 0.001, respectively) (**Table 2**). Even in adjusted models, the results remained the same after adjusting for potential covariates.

3.3 Subgroup and sensitivity analysis

Finally, we undertook two subgroup analyses of the study cohort (**Figure 2 and Figure 3**). In analyses stratified according to infants' sex, the use of ART was significantly associated with a higher social behaviour score compared with SC in male infants (adjusted β = 1.91; 95% CI 0.42–3.40; P = 0.012); however, no significant association was observed in female infants (adjusted β = 0.36; 95% CI -0.92–1.64; P = 0.584) (**Figure 2 and Supplementary Table S2**). When stratified according to plurality, the difference in the social behaviour score between the two groups was still persisted in singleton (adjusted β = 1.03; 95% CI 0.02–2.04; P = 0.046) **(Figure 3 and Supplementary Table S3**).

To eliminate the effect of preterm and LBW on neurodevelopmental outcomes, we excluded preterm (gestation age, <37 weeks) and LBW (birth weight, <2,500 g) infants from the sensitivity analysis. In the offspring with full term infant and normal birth weight, multivariate linear regression analysis showed similar results to those found in the overall study cohort (adjusted β = 1.07; 95% CI 0.04–2.09; P = 0.041, adjusted β = 1.17; 95% CI 0.16–2.17; P = 0.022, respectively) (**Figure 4 and Supplementary Table S4**).

4. DISCUSSION

In this population-based prospective longitudinal cohort study, offspring in the ART group

appeared to have similar motor skills, adaptive behaviour, language skills and total DQ scores to those in the SC group at age 12 months. However, social behaviour domain scores at age 12 months were slightly higher in those in the ART group compared with those in the SC group, even in the IPTW models. This small difference was more pronounced in males and in singleton.

Although ART treatment has resulted in the birth of millions worldwide, concerns have been raised regarding adverse outcomes on child development and wellbeing. The prevalence of ND among ART offspring identified in this study was 37 per 1000 live births, which was higher than that identified among SC offspring (29 per 1000). Clinical studies examining ART usage and subsequent offspring neurodevelopmental outcomes have yielded inconsistent results. Some studies have indicated an increased risk of neurodevelopmental disorders among ART-conceived offspring. For example, a Western Australian cohort study suggested that children who were conceived using ART had a 1.58-fold increased risk of intellectual disability (Hansen et al. 2018). Another Swedish prospective cohort study reported that ART was associated with a significantly increased risk of mental retardation (Sandin et al. 2013). However, our results showed that total neurodevelopment outcomes were not affected by the mode of conception. These findings were similar to a systematic review of current literature on neurodevelopmental outcomes in children conceived using ART, which concluded that ART has comparable outcomes overall when compared with those of children born after SC, especially in toddlers and preschool children (Bay et al. 2013). Another prospective cohort study used a different assessment method (Bayley scales) to determine neural outcomes and reported that children conceived using ART were also likely to have comparable neurodevelopmental outcomes (including cognitive, motor, and language skills) when compared with SC controls at age two years (Balayla et al. 2017). Moreover, one retrospective cohort study reported that preterm infants conceived using ART had lower rates of adverse neurodevelopmental outcomes at 18 to 24 months than preterm SC infants (Roychoudhury et al. 2021). These data suggest that exposure to ART does not lead to significant deleterious neurological outcome changes in offspring at a young age. The social behaviour score was slightly higher in ART offspring than in the controls, but the difference between two groups was small and not clinically important.

The increase in the use of ART has been reported to be related to rising rates of multiplets,

premature delivery, and LBW, as well as subsequent adverse obstetric and perinatal outcomes (Cavoretto et al. 2022, Nyboe Andersen et al. 2009, Qin et al. 2016). In the present study, the rate of multiple births was 1.0% among the controls and 25.6% among the ART births, and the proportions of LBW and preterm births among the ART offspring were approximately five and four times higher than those among SC offspring, respectively. It has been reported that infants born to infertile couples have a higher risk of multiplets, preterm birth, and LBW conditions (Goisis et al. 2019) and that these adverse perinatal outcomes were associated with neurodevelopmental impairment (Lin et al. 2020, Volpe 2009). After children with LBW or preterm were excluded, small differences of social behaviour skill still can be found between two groups. In addition, the relationship between better social behaviour skill and exposure to ART persists even in singleton.

Further highly developed stratified analyses have also shown that ART exposure was slightly associated with better social behaviour skill in males. Some previous studies have suggested that infants' sex may have an important influence on mediating the neurodevelopmental effects of prenatal exposure (Massey et al. 2018, Bale 2016). Male infants have been reported to be more vulnerable to early exposures than female infants (Benevenuto et al. 2017), which may explain why the difference in social behaviour skills was only found in male infants in our study.

To further verify the reliability of our results, we attempted to adjust for confounders through matching the two groups, which has not frequently been undertaken in previous related-cohort studies. Importantly, after adjusting for neurodevelopment-related confounders, no significant differences were found between the ART and control groups in the infants' development outcomes, except for social behaviour skills. But, two models unexpectedly showed that exposure to ART during pregnancy appeared to be adversely associated with the development of fine motor skills. Studies involving children conceived after ART have been frequently undertaken, with most studies predominantly focussing on risk factors in relation to total development (including gross and fine motor skills, and language and cognitive skills), and largely reporting no association between such risk factors and total development (Yeung et al. 2016). However, a Japanese birth cohort study reported that singleton females conceived after IVF had a higher risk of fine motor delay than those conceived spontaneously, even after adjusting for obstetric and perinatal factors

(Miyake et al. 2022). Therefore, further cohort studies with larger sample sizes may be required to verify our findings.

Finally, the results reflected no increased risks of gross motor skills, adaptive behaviour, and language skills in children born after ART compared with spontaneously conceived children. Although previous studies have shown that motor development is delayed by 16-18 months (Zhu et al. 2009). A prospective evidence of motor skills at 12 months old evaluated with standardized evaluation method is very rare in the literature. In our current cohort, we did not observe significant associations between exposure to ART and delayed gross motor skills of offspring in their early life. In addition, a controversy seems to be found in children's language development outcome after ART. An earlier study reported lower scores on receptive language skills in infants born after ART(Gibson et al. 1998). But another one found children concepted via ART was associated with decreased odds of language score of <85 using Bayley-III evaluation scale (aOR, 0.67; 95% CI, 0.50-0.88)(Roychoudhury et al. 2021). There is a call for prospective assessment of offspring's language skills by ART as we did in our study, in which we also find no significant delay in language development. In line with our findings, a recent prospective, longitudinal cohort in Canada carried out to assess the same question in this field, which found no significant difference in language and cognitive behaviour scores at 24 months of age(Balayla et al. 2017).

After all, with the increase of infertility year by year, the ART has become a rapidly evolving field. But it raises a number of concerns about potentially adverse consequences for child development. In particular, it has long been debated whether the use of ART is related to mental retardation. The most inspiring results are on neurodevelopmental outcomes in early childhood, with this prospective study reporting comparable gross motor skills, adaptive behaviour, language skills and total development levels between children born undergoing ART and spontaneously conceived controls. Moreover, consistent with our findings, two large meta-analyses had suggested the risk of adverse neurodevelopmental outcomes was not significantly different between offspring conceived vis ART and natural conception(Gibson et al. 1998, Bay et al. 2013). The replication of similar findings in above-mentioned studies is encouraging and may reassure couples using ART. Thus, our findings should be interpreted with caution that ART conceived infants are not worse than those conceived spontaneously, at least early in life. However, we

cannot preclude the possibility that some slight differences in cognitive, socio-emotional, and psychomotor development may appear later in life; therefore, we recommend a longer follow-up of children from prospective studies such as our study.

Our study had some limitations. First, we followed up the children for 12 months only, and these children may have been too young to undergo continuous neurodevelopmental monitoring and evaluation. However, validated evidence from studies concerning adolescents and young adults who were conceived using ART remains limited. These age groups are particularly important because impairments in complex neurodevelopmental and cognitive functioning may only emerge as children reach this later period of development. Conversely, early detection of changed performance may not accurately predict quality of performance later in life. Second, this cohort study was limited owing to its smaller sample size, such that the absolute incidence of ND was rare among the offspring. Moreover, given the relatively few offspring with developmental delay in the ART group, binary logistic regression analysis would have been more difficult to undertake, as it might have reduced the statistical power. Finally, although we adjusted for many confounders, there may have been other potential variables, such as children's thyroid functions and feeding status, that were not included. However, our study provides preliminary confirmation that ART offspring had comparable neurodevelopmental outcomes to SC controls. Further prospective well-designed studies with larger sample sizes, longer follow-up durations, and detailed records of reproductive history and pregnancy complications are recommended.

5. CONCLUSIONS

In summary, based on our analyses, no significant differences were found in terms of neurodevelopmental outcomes between ART-conceived and SC offspring after 12 months of follow-up. While social behaviour scores were slightly more elevated in the ART group, the difference between the two groups was small and not important clinically. Our findings may provide new information in evaluating the potential benefits and risks of ART. However, the absence of serious adverse neurodevelopmental outcomes found in this study does not imply that follow-ups and assessment after ART are not important.

Abbreviations

ART: Assisted reproductive technologies; SC: Spontaneous conceptions; BMI: Body mass index;

GDS: Gesell developmental scale; DQ: development quotient; ND: Neurodevelopmental delay; ICSI: Intracytoplasmic sperm injection; IVF: In vitro fertilization; LBW: Low birth weight; PSM: propensity score matching; IPTW: inverse probability of treatment weighting; CNBC: the China National Birth Cohort study; IPMCH: the International Peace Maternity and Child Health Hospital. GnRH: gonadotropin-releasing hormone. SD: standard deviation.

Acknowledgements

We are grateful to both the pregnant women and their children for participating in this study. Special thanks to the IPMCH Birth Cohort team for their willing participation and cooperation during the implementation of the survey.

Author contributions

Z.L. initiated, conceived, and supervised the study. W.L., J.Z., and Z.L. were involved in the study design. All authors were involved in data collection and long-term follow-up. W.L. and H.L. organised clinical information. W.L., M.N., J.X., X.W., and B.W. performed statistical analyses. W.L. prepared the manuscript. W.L. and Z.L. proofread the manuscript and take responsibility for the integrity of the data. All authors critically reviewed and provided feedback on the drafts and approved the final version.

Funding

This work was supported by the National Key R&D Program of China (2022YFC2702903), the National Natural Science Foundation of China Grants (81974232，82271742), the Clinical Research Plan of SHDC (SHDC2020CR6027，SHDC22022303), the Program of Shanghai Academic Research Leader (21XD1403700), the Interdisciplinary Program of Shanghai Jiao Tong University (YZ2021ZD29) and the Shanghai Municipal Science and Technology Major Project (20Z11900602).

Availability of data and materials

After publication, the data collected for the cohort (including data of deidentified participant) could be accessed on reasonable request to the corresponding author.

Declarations

Ethics approval and consent to participate

All methods and study protocols were approved by the Medical Ethics Committee of the

International Peace Maternity and Child Health Hospital, affiliated with Shanghai Jiao Tong University School of Medicine (approval number: GKLW-2016-21). Each participant provided informed consent in accordance with the Declaration of Helsinki.

Consent for publication

Not applicable.

Competing interests

There are no conflicts of interest to declare.

Declaration of Competing Interests

There are no conflicts of interest to declare.

References

Balayla, J., O. Sheehy, W. D. Fraser, J. R. Seguin, J. Trasler, P. Monnier, A. A. MacLeod, M. N. Simard, G. Muckle, A. Berard, D. S. R. G. F. t. I. R. N. i. P. o. Quebec & O. Eastern (2017) Neurodevelopmental Outcomes After Assisted Reproductive Technologies. *Obstet Gynecol,* 129**,** 265-272.

Bale, T. L. (2016) The placenta and neurodevelopment: sex differences in prenatal vulnerability. *Dialogues Clin Neurosci,* 18**,** 459-464.

Barbuscia, A., P. Martikainen, M. Myrskyla, H. Remes, E. Somigliana, R. Klemetti & A. Goisis (2020) Maternal age and risk of low birth weight and premature birth in children conceived through medically assisted reproduction. Evidence from Finnish population registers. *Hum Reprod,* 35**,** 212-220.

Bay, B., E. L. Mortensen & U. S. Kesmodel (2013) Assisted reproduction and child neurodevelopmental outcomes: a systematic review. *Fertil Steril,* 100**,** 844-53.

Benevenuto, S. G., M. D. Domenico, M. A. G. Martins, N. S. Costa, A. R. L. de Souza, J. L. Costa, M. F. M. Tavares, M. Dolhnikoff & M. M. Veras (2017) Recreational use of marijuana during pregnancy and negative gestational and fetal outcomes: An experimental study in mice. *Toxicology,* 376**,** 94-101.

Blencowe, H., S. Cousens, D. Chou, M. Oestergaard, L. Say, A. B. Moller, M. Kinney, J. Lawn & B. T. S. P. B. Action (2013) Born Too Soon: The global epidemiology of 15 million preterm births. *Reproductive Health,* 10.

Boulet, S. L., R. S. Kirby, J. Reefhuis, Y. Zhang, S. Sunderam, B. Cohen, D. Bernson, G. Copeland, M. A. Bailey, D. J. Jamieson, D. M. Kissin & C. States Monitoring Assisted Reproductive Technology (2016) Assisted Reproductive Technology and Birth Defects Among Liveborn Infants in Florida, Massachusetts, and Michigan, 2000-2010. *JAMA Pediatr,* 170**,** e154934.

Cavoretto, P. I., V. Giorgione, A. Sotiriadis, P. Vigano, E. Papaleo, A. Galdini, G. Gaeta & M. Candiani (2022) IVF/ICSI treatment and the risk of iatrogenic preterm birth in singleton pregnancies: systematic review and meta-analysis of cohort studies. *J Matern Fetal Neonatal Med,* 35**,**

1987-1996.

Cui, L., W. Zhou, B. Xi, J. Ma, J. Hu, M. Fang, K. Hu, Y. Qin, L. You, Y. Cao, L. Yang, L. Yang, C. Ma, W. Shui, M. Wang, M. Zhao, J. Zhang & Z. J. Chen (2020) Increased risk of metabolic dysfunction in children conceived by assisted reproductive technology. *Diabetologia,* 63**,** 2150-2157.

Declercq, E., B. Luke, C. Belanoff, H. Cabral, H. Diop, D. Gopal, L. Hoang, M. Kotelchuck, J. E. Stern & M. D. Hornstein (2015) Perinatal outcomes associated with assisted reproductive technology: the Massachusetts Outcomes Study of Assisted Reproductive Technologies (MOSART). *Fertil Steril,* 103**,** 888-95.

Donald, K. A., C. J. Wedderburn, W. Barnett, R. T. Nhapi, A. M. Rehman, J. A. M. Stadler, N. Hoffman, N. Koen, H. J. Zar & D. J. Stein (2019) Risk and protective factors for child development: An observational South African birth cohort. *PLoS Med,* 16**,** e1002920.

Dror, R., G. Malinger, L. Ben-Sira, D. Lev, C. G. Pick & T. Lerman-Sagie (2009) Developmental outcome of children with enlargement of the cisterna magna identified in utero. *J Child Neurol,* 24**,** 1486-92.

Gibson, F. L., J. A. Ungerer, G. I. Leslie, D. M. Saunders & C. C. Tennant (1998) Development, behaviour and temperament: a prospective study of infants conceived through in-vitro fertilization. *Human Reproduction,* 13**,** 1727-1732.

Goisis, A., H. Remes, P. Martikainen, R. Klemetti & M. Myrskyla (2019) Medically assisted reproduction and birth outcomes: a within-family analysis using Finnish population registers. *Lancet,* 393**,** 1225-1232.

Guo, X. Y., X. M. Liu, L. Jin, T. T. Wang, K. Ullah, J. Z. Sheng & H. F. Huang (2017) Cardiovascular and metabolic profiles of offspring conceived by assisted reproductive technologies: a systematic review and meta-analysis. *Fertil Steril,* 107**,** 622-631 e5.

Halpert, J. (2012) Altering the primal environment: health effects associated with assisted reproductive technologies. *Environ Health Perspect,* 120**,** a390-5.

Hansen, M., K. R. Greenop, J. Bourke, G. Baynam, R. J. Hart & H. Leonard (2018) Intellectual Disability in Children Conceived Using Assisted Reproductive Technology. *Pediatrics,* 142.

Hargreave, M., A. Jensen, M. K. Hansen, C. Dehlendorff, J. F. Winther, K. Schmiegelow & S. K. Kjaer (2019) Association Between Fertility Treatment and Cancer Risk in Children. *Jama-Journal of the American Medical Association,* 322**,** 2203-2210.

Lin, D. X., P. S. Li, D. Z. Fan, G. D. Chen, S. Z. Wu, S. X. Ye, H. T. Ma, J. M. Rao, Z. X. Zhou, M. Zeng, Z. Huang, X. L. Guo & Z. P. Liu (2021) Association between IVF/ICSI treatment and preterm birth and major perinatal outcomes among dichorionic-diamnionic twin pregnancies: A seven-year retrospective cohort study. *Acta Obstetricia Et Gynecologica Scandinavica,* 100**,** 162-169.

Lin, L., E. Amissah, G. D. Gamble, C. A. Crowther & J. E. Harding (2020) Impact of macronutrient supplements on later growth of children born preterm or small for gestational age: A systematic review and meta-analysis of randomised and quasirandomised controlled trials. *PLoS Med,* 17**,** e1003122.

Lv, H., F. Diao, J. Du, T. Chen, Q. Meng, X. Ling, H. Li, C. Song, Q. Xi, Y. Jiang, Y. Xu, S. Tao, L. Huang, M. Wen, M. Peng, C. Liu, Q. Lu, Y. He, Y. Yin, X. Liu, B. Xu, X. Han, K. Zhou, T. Jiang, Y. Zhao, H. Ma, G. Jin, Y. Xia, J. Liu, Y. Lin, Z. Hu & H. Shen (2021) Assisted reproductive technology and birth defects in a Chinese birth cohort study. *Lancet Reg Health West Pac,* 7**,** 100090.

Massey, S. H., D. K. Mroczek, D. Reiss, E. S. Miller, J. A. Jakubowski, E. K. Graham, S. M. Shisler, M. McCallum, M. A. Huestis, J. M. Ganiban, D. S. Shaw, L. D. Leve, R. D. Eiden, L. R. Stroud & J. M. Neiderhiser (2018) Additive drug-specific and sex-specific risks associated with co-use of marijuana and tobacco during pregnancy: Evidence from 3 recent developmental cohorts (2003-2015). *Neurotoxicol Teratol,* 68**,** 97-106.

Miyake, T., M. Yamamoto, K. Sakurai, A. Eguchi, M. Yoshida, C. Mori & J. E. C. S. JEC (2022) Neurological development in 36-month-old children conceived via assisted reproductive technology: The Japan Environment and Children's Study. *Reproductive Medicine and Biology,* 21.

Nyboe Andersen, A., V. Goossens, S. Bhattacharya, A. P. Ferraretti, M. S. Kupka, J. de Mouzon, K. G. Nygren, f. t. E. S. o. H. R. European Ivf-monitoring Consortium & Embryology (2009) Assisted reproductive technology and intrauterine inseminations in Europe, 2005: results generated from European registers by ESHRE: ESHRE. The European IVF Monitoring Programme (EIM), for the European Society of Human Reproduction and Embryology (ESHRE). *Hum Reprod,* 24**,** 1267-87.

Pontesilli, M., M. H. Hof, A. C. J. Ravelli, A. J. van Altena, A. T. Soufan, B. W. Mol, E. H. Kostelijk, E. Slappendel, D. Consten, A. E. P. Cantineau, L. A. J. van der Westerlaken, W. van Inzen, J. C. M. Dumoulin, L. Ramos, E. B. Baart, F. J. M. Broekmans, P. M. Rijnders, M. H. J. M. Curfs, S. Mastenbroek, S. Repping, T. J. Roseboom & R. C. Painter (2021) Effect of parental and ART treatment characteristics on perinatal outcomes. *Human Reproduction,* 36**,** 1640-1665.

Qin, J., X. Liu, X. Sheng, H. Wang & S. Gao (2016) Assisted reproductive technology and the risk of pregnancy-related complications and adverse pregnancy outcomes in singleton pregnancies: a meta-analysis of cohort studies. *Fertil Steril,* 105**,** 73-85 e1-6.

Roychoudhury, S., A. Lodha, A. Synnes, A. Abou Mehrem, R. Canning, R. Banihani, M. Beltempo, K. Theriault, J. M. Yang, P. S. Shah, A. S. Soraisham, B. C. N. Network, C. P. B. Network & C. N. F.-U. Networ (2021) Neurodevelopmental outcomes of preterm infants conceived by assisted reproductive technology. *American Journal of Obstetrics and Gynecology,* 225.

Sandin, S., K. G. Nygren, A. Iliadou, C. M. Hultman & A. Reichenberg (2013) Autism and Mental Retardation Among Offspring Born After In Vitro Fertilization. *Jama-Journal of the American Medical Association,* 310**,** 75-84.

Scharf, R. J., A. Stroustrup, M. R. Conaway & M. D. DeBoer (2016) Growth and development in children born very low birthweight. *Archives of Disease in Childhood-Fetal and Neonatal Edition,* 101**,** F433-F438.

Volpe, J. J. (2009) Brain injury in premature infants: a complex amalgam of destructive and

developmental disturbances. *Lancet Neurol,* 8**,** 110-24.

Wang, A. Y., A. A. Chughtai, K. Lui & E. A. Sullivan (2017) Morbidity and mortality among very preterm singletons following fertility treatment in Australia and New Zealand, a population cohort study. *Bmc Pregnancy and Childbirth,* 17.

Wang, C., H. Lv, X. Ling, H. Li, F. Diao, J. Dai, J. Du, T. Chen, Q. Xi, Y. Zhao, K. Zhou, B. Xu, X. Han, X. Liu, M. Peng, C. Chen, S. Tao, L. Huang, C. Liu, M. Wen, Y. Jiang, T. Jiang, C. Lu, W. Wu, D. Wu, M. Chen, Y. Lin, X. Guo, R. Huo, J. Liu, H. Ma, G. Jin, Y. Xia, J. Sha, H. Shen & Z. Hu (2021) Association of assisted reproductive technology, germline de novo mutations and congenital heart defects in a prospective birth cohort study. *Cell Res,* 31**,** 919-928.

Wang, C., Q. Wang, B. Xiang, S. Chen, F. Xiong & Y. Ji (2018) Effects of Propranolol on Neurodevelopmental Outcomes in Patients with Infantile Hemangioma: A Case-Control Study. *Biomed Res Int,* 2018**,** 5821369.

Yang, X., Y. Li, C. Li & W. Zhang (2014) Current overview of pregnancy complications and live-birth outcome of assisted reproductive technology in mainland China. *Fertil Steril,* 101**,** 385-91.

Yeung, E. H., R. Sundaram, E. M. Bell, C. Druschel, C. Kus, A. Ghassabian, S. Bello, Y. L. Xie & G. M. B. Louis (2016) Examining Infertility Treatment and Early Childhood Development in the Upstate KIDS Study. *Jama Pediatrics,* 170**,** 251-258.

Zhang, L., W. Zhang, H. Xu & K. Liu (2021) Birth defects surveillance after assisted reproductive technology in Beijing: a whole of population-based cohort study. *BMJ Open,* 11**,** e044385.

Zhu, J. L., O. Basso, C. Obel, D. Hvidtjorn & J. Olsen (2009) Infertility, infertility treatment and psychomotor development: the Danish National Birth Cohort. *Paediatric and Perinatal Epidemiology,* 23**,** 98-106.

Author Biography

Zhiwei Liu, Departments of Neonatology, International Peace Maternity and Child Hospital, School of Medicine, Shanghai Jiao Tong University, 910# Hengshan Road, Shanghai 200030, China. Fax: 86-21-64073899, Tel: 86-21-64073899, E-mail: liuzhiwei@hotmail.com

Figure titles and legends

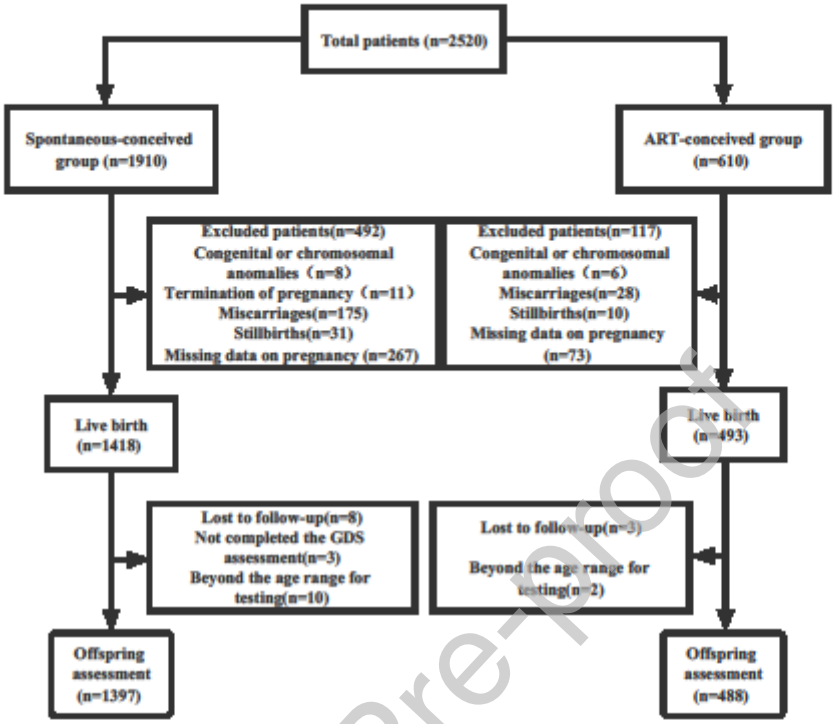

Figure 1. Flow diagram of study participants.

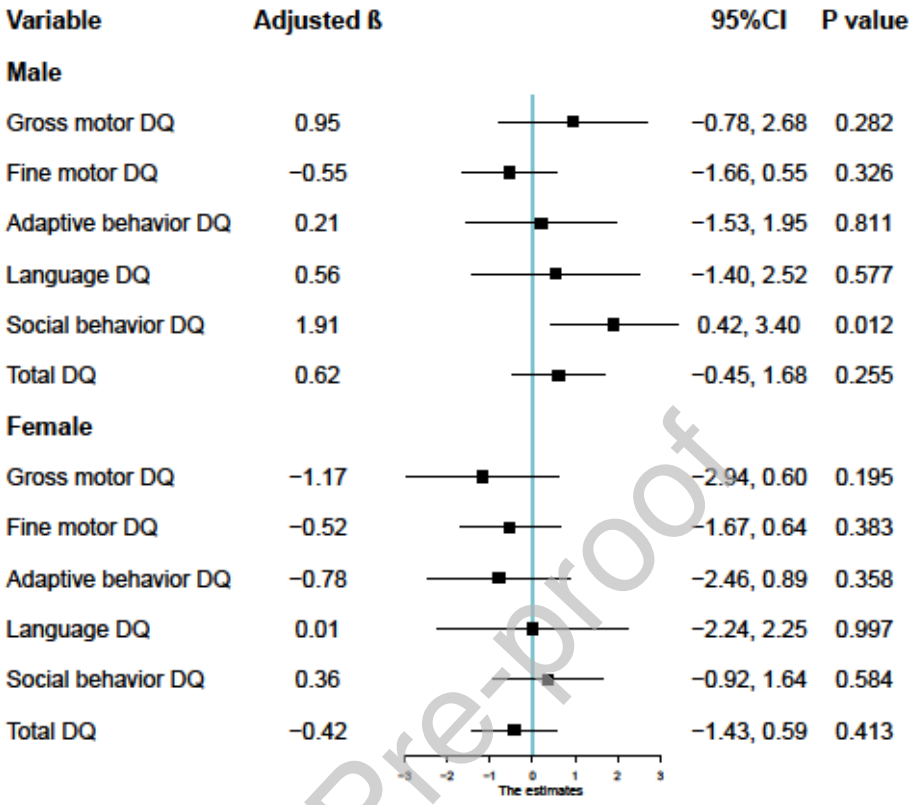

Figure 2. Subgroup analysis: Linear regressions of neurodevelopmental scores among ART relative to spontaneous conception stratified by infants' sex. DQ, development quotient; , , regression coefficient; CI, confidence interval. The adjusted coefficient and 95% CI were calculated from the multivariable linear regression.

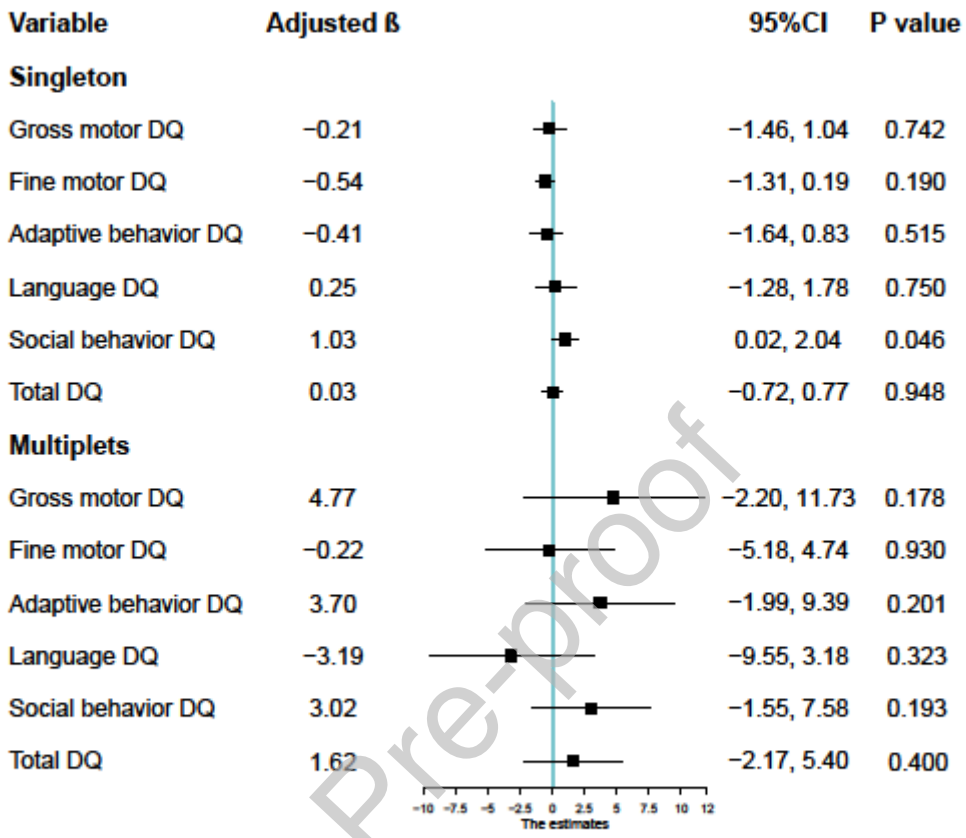

Figure 3. Subgroup analysis: Linear Regressions of Neurodevelopmental Scores Among ART Relative to Spontaneous Conception stratified by plurality. DQ, development quotient; ¸ , regression coefficient; CI, confidence interval. The adjusted¸ coefficient and 95% CI were calculated from the multivariable linear regression.

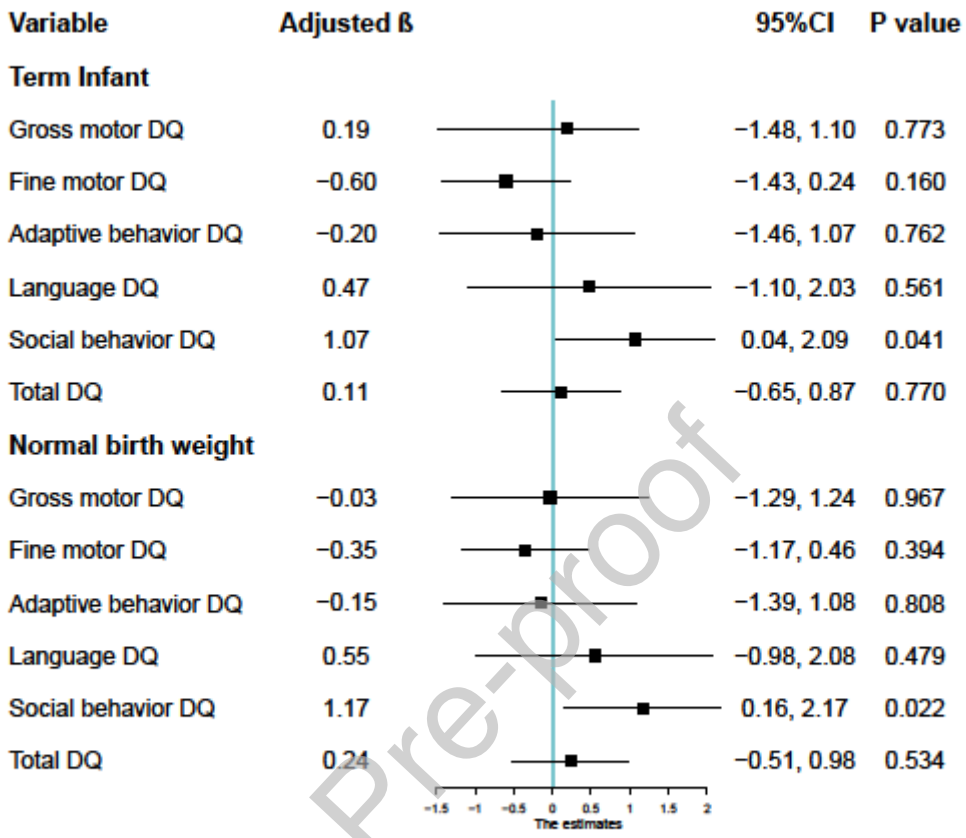

Figure 4. Sensitivity analysis: Linear regressions of neurodevelopmental scores among ART relative to spontaneous conception. Term infant means gestational age longer than 37 weeks. Normal birth weight means birth weight higher than 2500g. DQ, development quotient; ß, regression coefficient; CI, confidence interval. The adjusted ß coefficient and 95% CI were calculated from the multivariable linear regression.

Assisted reproductive technology and neurodevelopmental outcomes in offspring: a prospective birth cohort study in China

Table 1
Basic characteristics of the study population

	Full Population, N (%)			Weighted Population, IPTW, N (%)		
	SC (N = 1397)	ART (N = 488)	*P* value	SC (N = 1844)	ART (N = 1900)	*P* value
[a]**Parental characteristics**						
Maternal age, year, No. (%)			<0.001			0.578
<35	1090(78.0)	317(65.0)		1395(75.7)	1411(74.3)	
≥35	307 (22.0)	171(35.0)		448(24.3)	489(25.7)	
Paternal age, year, No. (%)			<0.001			0.802
<35	971(69.5)	249(52.6)		1197(64.9)	1219(57.9)	
≥35	426(30.5)	224(47.4)		647(35.1)	681(42.1)	
Plurality, No. (%)			<0.001			0.268
Singleton	1383(99.0)	363(74.4)		1748(94.8)	1762(92.7)	
Multiplets	14(1.0)	125(25.6)		96(5.2)	138(7.3)	
Maternal pre-pregnancy BMI, kg/m2, No. (%)			<0.001			0.716
<18.5	185 (13.3)	41(8.8)	<0.001	224(12.1)	262(7.8)	
18.5-24.5	982(70.3)	316(67.7)		1304(70.7)	1298(68.3)	
>=24.5	229(16.4)	110(23.6)		316(17.1)	341(17.9)	
Household income, CNY, No. (%)			<.001			0.734
<100,000	75(5.4)	51(10.8)		118(6.4)	123(6.5)	
100,000–300,000	828(59.3)	283(59.8)		1081(58.7)	1155(60.8)	
300,000	494(35.4)	139(29.4)		644(34.9)	622(32.7)	
Maternal education, year, No. (%)			0.002			0.450
High school or lower	340(24.3)	150(31.7)		474(25.7)	530(27.9)	
College degree or above	1057(75.7)	323(68.3)		1369(74.3)	1370(72.1)	
Paternal education, year, No. (%)			0.002			0.790
High school or lower	317 22.7	141(29.8)		461(25.0)	490(25.8)	
College degree or above	1080(77.3)	332(70.2)		1382(75.0)	1410(74.2)	
Tobacco use during pregnancy, No. (%)	0 (0.0)	4(0.8)	0.004	0.0 (0.0)	17 (0.9)	0.110
Alcohol intake during pregnancy, No. (%)	45(3.2)	13(2.7)	0.607	4 (3.5)	55 (3.0)	0.709
Diseases during pregnancy						
[b]Diabetes, No. (%)	177(12.7)	90(18.6)	0.001	272(14.8)	378(20.1)	0.053
[c]Hypertension,No.(%)	61(4.4)	29(6.0)	0.151	104 (5.6)	105 (5.6)	0.969
Thyroid disease, No. (%)	177(12.7)	75(15.5)	0.118	235(12.8)	263(14.0)	0.607
Preeclampsia, No. (%)	30(2.2)	33(6.8)	<0.001	49 (2.7)	86 (4.6)	0.067
[a]**Newborn characteristics**						
Infants' sex, NO (%)			0.808			0.709

Male	707(50.6)	243(49.8)		927(50.3)	931(49.0)	
Female	690(49.4)	245(50.2)		917(49.7)	969(51.0)	
Birth weight, No. (%)			<0.001			0.972
≥2500	1357(97.1)	419(86.0)		1749(94.8)	1801(94.8)	
2500	40(2.9)	68(14.0)		95(5.2)	99(5.2)	
Gestational age, weeks, No. (%)			<0.001			0.659
<34	14 (1.0)	19(3.9)		48(2.6)	34(1.8)	
34-36+6	62 (4.4)	80(16.4)		120(6.5)	137(7.2)	
37 or above	1321 (94.6)	388(79.7)		1676(90.9)	1729(91.0)	
Cesarean section	576(41.2)	356(73.0)	<0.001	895(48.6)	925(48.7)	0.971

Abbreviation: IPTW, inverse probability of treatment weight; SC, Spontaneous-conception; ART, assisted reproductive technology; BMI, body mass index; CNY, China yuan.

a No missing for maternal age, plurality, infant sex, mode of delivery and neurodevelopment outcome. Other missing data are less than 2%.

b Includes chronic and gestational diabetes.

c Includes chronic and pregnancy-induced hypertension.

P values were calculated by Chi-squared test or Fisher's exact test for categorical variables and independent Student t-test or Mann Whitney U test (non-normally distributed variables) for continuous variables.

Data are number (percent).

Table 2
Characteristics of neurodevelopmental outcomes in offspring

Neurodevelopment outcome	Full Population			Weighted Population, IPTW			Matched Population, PSM		
	SC (N = 1397)	ART (N= 488)	*P* value	SC (N = 1850)	ART (N = 1897)	*P* value	SC (N = 1850)	ART (N = 1897)	*P* value
Gross motor DQ	105.25(10.32)	104.98(10.66)	0.646	105.22(10.38)	104.67(10.24)	0.098	104.88(10.35)	104.65(10.33)	0.203
Fine motor DQ	97.01 (6.66)	96.56 (7.06)	0.171	96.94(6.61)	96.33(6.22)	**0.002**	96.97(6.58)	96.56(6.59)	**<0.001**
Adaptive behaviour DQ	94.57(10.29)	93.99(10.01)	0.300	94.36(10.29)	94.46(9.66)	0.982	94.22(10.41)	94.22(10.02)	0.209
Language DQ	86.90(12.90)	86.25(12.14)	0.702	86.73(12.77)	86.81(12.27)	0.301	86.70(12.83)	86.8(12.47)	0.969
Social behaviour DQ	100.32 (8.44)	100.96 (7.99)	0.211	100.13(8.52)	101.09(7.56)	0.005	100.20(8.38)	101.10(7.84)	0.572
Total DQ	96.81 (6.24)	96.55 (6.25)	0.423	96.68(6.23)	96.41(5.82)	0.889	96.6(6.25)	96.67(6.13)	0.323

Abbreviation: SC, Spontaneous-conception; ART, assisted reproductive technology; DQ, development quotient.
Data are mean (standard deviation, SD)

Table 3

Linear regressions of neurodevelopmental scores among ART relative to spontaneous conception

Variable	Crude Model			Adjusted Model		
	B1(SE)	95%CI	*P*	B1(SE)	95%CI	*P*
Full Population						
Gross motor DQ	-0.28(0.55)	-1.35, 0.79	0.610	-0.11(0.63)	-1.34, 1.13	0.864
Fine motor DQ	-0.46(0.36)	-1.16, 0.24	0.198	-0.56(0.41)	-1.36, 0.24	0.170
Adaptive behavior DQ	-0.58(0.54)	-1.63, 0.48	0.282	-0.29(0.61)	-1.49, 0.91	0.637
Language DQ	-0.65(0.69)	-1.96, 0.66	0.330	0.21(0.76)	-1.27, 1.70	0.778
Social behavior DQ	0.64(0.44)	-0.22, 1.50	0.143	1.08(0.50)	0.10, 2.06	**0.031**
Total DQ	-0.27(0.33)	-0.91, 0.38	0.419	0.07(0.37)	-0.66, 0.80	0.856
IPTW						
Gross motor DQ	-0.55 (0.48)	-1.48, 0.39	0.250	-0.35 (0.48)	-1.29, 0.60	0.473
Fine motor DQ	-0.61 (0.30)	-1.19, -0.03	***0.041***	-0.72 (0.30)	-1.30, -1.42	***0.015***
Adaptive behavior DQ	0.10 (0.46)	-0.80, 1.00	0.827	-0.11 (0.46)	-1.01, 0.79	0.810
Language DQ	0.08 (0.58)	-1.05, 1.21	0.893	0.00 (0.58)	-1.13, 1.13	0.998
Social behavior DQ	0.96 (0.37)	0.23, 1.69	***0.010***	0.99 (0.37)	0.26, 1.71	***0.008***
Total DQ	-0.00 (0.28)	-0.55, 0.54	0.992	0.04 (0.28)	-0.58, 0.50	0.891
PSM						
Gross motor DQ	-0.66 (0.67)	-1.98, 0.67	0.330	-0.36 (0.69)	-1.70, 1.00	0.606
Fine motor DQ	-0.99 (0.43)	-1.83, -0.14	***0.023***	-1.05 (0.43)	-1.90, -0.20	***0.016***
Adaptive behavior DQ	-0.83 (0.66)	-2.11, 0.46	0.207	-0.98 (0.66)	-2.26, 0.31	0.137
Language DQ	-0.09 (0.84)	-1.73, 1.56	0.917	0.08 (0.84)	-1.57, 1.72	0.928
Social behavior DQ	0.75 (0.59)	-0.40, 1.89	0.202	0.88 (0.59)	-0.28, 2.04	0.136
Total DQ	-0.36 (0.41)	-1.17, 0.45	0.380	0.29(0.41)	-1.09, 0.52	0.488

Abbreviation: IPTW, Inverse probability of treatment weighting; PSM, Propensity score matching; DQ, development quotient; B1, linear regression β coefficient; SE, standard error; CI, confidence interval. Whereas B1 represents the change in the score units of each scale for ART conception relative to a spontaneous conception.

Spontaneous conception group as a reference.

Statistical test used: crude and adjusted linear regression.

Adjusted Model, adjusted for father age, maternal age, maternal education, income, maternal pre-pregnancy BMI, smoking intake, alcohol consumption during pregnancy, plurality, and infants' sex.