ORIGINAL ARTICLE: ENVIRONMENT AND EPIDEMIOLOGY

# Assisted reproductive technology and perinatal outcomes: conventional versus discordant-sibling design

Nafeesa N. Dhalwani, Ph.D.,[a,b,c] Sheree L. Boulet, Dr.P.H.,[a] Dmitry M. Kissin, M.D.,[a] Yujia Zhang, Ph.D.,[a] Patricia McKane, M.P.H.,[d] Marie A. Bailey, M.S.W.,[e] Maria-Elena Hood, M.P.H.,[f] and Laila J. Tata, Ph.D.[b]

[a] Division of Reproductive Health, Centers for Disease Control and Prevention, Atlanta, Georgia; [b] Division of Epidemiology and Public Health, University of Nottingham, Nottingham, United Kingdom; [c] Diabetes Research Centre, University of Leicester, Leicester General Hospital, Leicester, United Kingdom; [d] Michigan Department of Health and Human Services, Lansing, Michigan; [e] Florida Department of Health, Tallahassee, Florida; and [f] Massachusetts Department of Public Health, Boston, Massachusetts

**Objective:** To compare risks of adverse perinatal outcomes between assisted reproductive technology (ART) and naturally conceived singleton births using a dual design approach.
**Design:** Discordant-sibling and conventional cross-sectional general population comparison.
**Setting:** Not applicable.
**Patient(s):** All singleton live births, conceived naturally or via ART.
**Intervention(s):** None.
**Main Outcome Measure(s):** Birth weight, gestational age, low birth weight, preterm delivery, small for gestational age (SGA), low Apgar score.
**Result(s):** A total of 32,762 (0.8%) of 3,896,242 singleton live births in the three states were conceived via ART. In 6,458 sibling pairs, ART-conceived singletons were 33 g lighter (adjusted $\beta = -33.40$, 95% confidence interval [CI], −48.60, −18.21) and born half a day sooner ($\beta = -0.58$, 95% CI, −1.02, −0.14) than singletons conceived naturally. The absolute risk of low birth weight and preterm birth was 6.8% and 9.7%, respectively, in the ART group and 4.9% and 7.9%, respectively, in the non-ART group. The odds of low birth weight were 33% higher (adjusted odds ratio [aOR] = 1.33; 95% CI, 1.13, 1.56) and 20% higher for preterm birth (aOR = 1.20; 95% CI, 1.07, 1.34). The odds of SGA and low Apgar score were not significantly different in both groups (aOR = 1.22; 95% CI, 0.88, 1.68; and aOR = 0.75; 95% CI, 0.54, 1.05, respectively). Results of conventional analyses were similar, although the magnitude of risk was higher for preterm birth (aOR, 1.51; 95% CI 1.46, 1.56).
**Conclusion(s):** Despite some inflated risks in the general population comparison, ART remained associated with an increased likelihood of low birth weight and preterm birth when underlying maternal factors were kept constant using discordant-sibling comparison. (Fertil Steril® 2016;■:■–■. ©2016 by American Society for Reproductive Medicine.)
**Key Words:** Assisted reproductive technology, low birth weight, preterm birth, small for gestational age, discordant-sibling design

**Discuss:** You can discuss this article with its authors and with other ASRM members at http://fertstertforum.com/dhalwanin-art-perinatal-outcomes/

Use your smartphone to scan this QR code and connect to the discussion forum for this article now.*

* Download a free QR code scanner by searching for "QR scanner" in your smartphone's app store or app marketplace.

Received December 4, 2015; revised April 11, 2016; accepted April 25, 2016.
N.N.D. has nothing to disclose. S.L.B. has nothing to disclose. D.M.K. has nothing to disclose. Y.Z. has nothing to disclose. P.M. has nothing to disclose. M.A.B. has nothing to disclose. M.-E.H. has nothing to disclose. L.J.T. has nothing to disclose.
Supported by the Building Experience and Skills Travel Scholarship awarded by the Graduate School, University of Nottingham for the duration of the research project (to N.N.D.); the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care—East Midlands (NIHR CLAHRC-EM; to N.N.D.); the Leicester Clinical Trials Unit (to N.N.D.); and the NIHR Leicester-Loughborough Diet, Lifestyle and Physical Activity Biomedical Research Unit, which is a partnership among University Hospitals of Leicester National Health Service (NHS) Trust, Loughborough University, and the University of Leicester (to N.N.D.).
The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention, the NHS, the NIHR, or the Department of Health.
Reprint requests: Nafeesa N. Dhalwani, Ph.D., Diabetes Research Centre, University of Leicester, Leicester General Hospital, Gwendolen Road, Leicester, Leicestershire LE54PW, United Kingdom (E-mail: nnd2@le.ac.uk).

Fertility and Sterility® Vol. ■, No. ■, ■ 2016 0015-0282/$36.00

http://dx.doi.org/10.1016/j.fertnstert.2016.04.038

Reporting of fertility problems has increased substantially over the last decades, with more couples seeking medical consultation as a result of delaying pregnancy and increased awareness and wider availability of fertility treatments (1). Consequently, the use of assisted reproductive technology (ART) has increased such that ART now accounts for approximately 1.5% of live births in the United States and 2% in the United Kingdom (UK) (2, 3). The association between ART and adverse perinatal outcomes has previously been assessed for different types of ART, causes of infertility, and singleton and twin or higher order deliveries (4–8). A population-based study from the United States linking the ART surveillance system data with Massachusetts live birth–infant death records data for 1997–98 found the risk of preterm delivery (<37 weeks of gestation) and low birth weight (birth weight <2,500 g) to be over twice as high in singletons conceived through ART compared with natural births (5). A more recent study of 306,995 births in South Australia showed that compared with non-ART singletons, ART singletons had compromised perinatal outcomes varying by the type of ART used (9). One potential explanation for ART-associated adverse perinatal outcomes is underlying maternal or paternal factors that may result in infertility and also cause adverse perinatal outcomes. However, studies often compare couples undergoing ART with couples conceiving naturally, and most have not been able to adjust for underlying parental factors that may result in infertility and consequently adverse perinatal outcomes. One alternative approach is to compare pregnancies in the same women using discordant-sibling analysis, which keeps many maternal factors fairly constant between comparisons, minimizing related residual confounding. Only two studies to date have used this approach and found conflicting results on the association between ART and perinatal outcomes (10, 11).

Using the recent linkage of the National ART Surveillance System (NASS) with birth records from three U.S. states, we compared the risk of adverse perinatal outcomes in singleton infants conceived through ART with those conceived naturally using a sibling analysis approach and compared results with those obtained with a conventional cross-sectional approach using all non-ART singletons from the general population as the comparison group.

# METHODS

## Data Source

We used data from the States Monitoring Assisted Reproductive Technology (SMART) Collaborative, a collaboration between the Centers for Disease Control and Prevention's (CDC) Division of Reproductive Health, Florida Department of Health, the Massachusetts Department of Public Health, and the Michigan Department of Community Health. As part of the SMART Collaborative, data from NASS, collected by CDC on all ART cycles performed in the United States, have been linked to birth certificate data from the three participating states. Data are linked using the Link Plus software (CDC) and a probabilistic linkage algorithm using maternal date of birth, infant date of birth, plurality, maternal residence ZIP code, and gravidity as the primary linkage variables, with a linkage rate of 90.2%. A detailed description of data collection and linkages for the SMART collaborative has been reported elsewhere (12, 13).

## Study Population

The study population included all singleton live births in the three participating states between 2000 and 2010. For Massachusetts and Florida, linked data were available for all study years, while for Michigan, linked data were only available until 2009. Infants born to mothers less than 20 years old at delivery and those having gestational age of less than 20 weeks (estimated clinically) or greater than 46 weeks and birth weight above or below the national reference range for gestational age (14) were excluded. ART births were identified using the NASS-linked data, which provided detailed information on ART procedures and reasons for infertility treatment.

To create discordant-sibling pairs we restricted the study population to singleton live births where one sibling was conceived through ART and the other was conceived without ART, regardless of the order of conception. In the instance of a woman having more than one ART or non-ART singleton live birth during the study period, one birth of each type was randomly selected to create the sibling pair.

## Main Outcome Measures

We examined birth weight as a continuous variable and also assessed whether the infant had low birth weight, defined as less than 2,500 g according to the World Health Organization (WHO) definition (15). Similarly, gestational age at delivery was used as a continuous variable, and we also assessed preterm birth, defined as birth before 37 weeks of gestation in line with the WHO definition (16). Small for gestational age (SGA) was defined as a birth weight below 2 SD of mean birth weight for gestational age according to the infant's sex in line with the International Societies of Pediatric Endocrinology and the Growth Hormone Research Society definition (17). We used this definition as opposed to using the more common definition (weight below 10th percentile for gestational age) as it identifies most infants at risk of developing childhood or adulthood morbidity (17) and also facilitates comparisons with the previous study (10). Apgar score is a scoring system to assess the clinical status of the newborn at 1 and 4 minutes and is composed of five components: heart rate, respiratory effort, muscle tone, reflex irritability, and color, each of which is given a score of 0, 1, or 2. A low Apgar score at 5 minutes has been shown to be associated with neonatal mortality and adverse neurological outcomes in infants (18). We categorized the 5-minute Apgar into normal ($\geq$7) and low (<7) in line with previous studies (5).

## Statistical Analysis

We compared baseline birth characteristics in the ART and non-ART groups using $\chi^2$-tests and assessed types of ART and reasons for treatment in the population using proportions. We used two approaches to multivariable analysis: one using births from the three states in a conventional