Comparison of birth weights in patients randomly assigned to fresh or frozen-thawed embryo transfer

Bruce S. Shapiro, M.D., Ph.D., Said T. Daneshmand, M.D., Carrie E. Bedient, M.D., and Forest C. Garner, M.Sc.
Fertility Center of Las Vegas and University of Nevada School of Medicine, Las Vegas, Nevada

Objective: To estimate birth weight differences between patients randomized to fresh or thawed ET.

Design: Post hoc analysis of results from two similar randomized trials.

Setting: Private fertility center.

Patient(s): One hundred thirty-four first-time IVF patients, ages 18–40 years at oocyte retrieval, who had live birth.

Intervention(s): Patients were randomly assigned to have either fresh blastocyst transfer or all bipronuclear oocytes frozen followed by thaw, extended culture, and blastocyst transfer in a subsequent cycle. Preimplantation genetic screening was not allowed.

Main Outcome Measure(s): Mean birth weight.

Result(s): After allowing for the contributions of multiple significant variables (gestational age at birth, the presence of a vanished twin, number of infants delivered) in multiple linear regression, the adjusted mean birth weight was 166 g (95% confidence interval, 43–290 g) lower after fresh blastocyst transfer when compared with transfer of blastocysts derived from thawed bipronuclear oocytes.

Conclusion(s): Birth weights are lower in cycles with fresh blastocyst transfer after controlled ovarian stimulation than in transfers of frozen–thawed embryos in the absence of ovarian stimulation. This finding confirms similar results reported in many retrospective studies.

Clinical Trial Registration Numbers: NCT00963625 and NCT00963079. (Fertil Steril® 2016;106:317–21. ©2016 by American Society for Reproductive Medicine.)

Key Words: Assisted reproduction, ovarian stimulation, embryo cryopreservation, birth weight

Discuss: You can discuss this article with its authors and with other ASRM members at http://fertstertforum.com/shapirobirthweights-comparison-fresh-frozen/
Nordic registries (20), 156 g in the United States (21), 91–100 g in Japan (22, 23), 145 g in Australia and New Zealand (24), and 80 g among singletons in Latin America (25). Multiple retrospective clinical studies have also found greater birth weight with frozen-thawed ET (FET) than with fresh transfer, with reported birth weight differences ranging from 50 to 218 g (26–29).

One potential cause of birth weight differences is a suboptimal uterine environment after COS exposure (21). Alternatively, cryopreservation might alter birth weight through embryonic effects. However, the contrasting patterns in birth weight effects after autologous fresh, autologous FET, fresh cycles of oocyte donation, and FET cycles using embryos derived from donor oocytes have been used to examine the potential for embryonic or uterine effects of cryopreservation. An American registry study reported no significant difference between the incidence of low birth weight after fresh donor and donor FET cycles, while finding increased incidence of low birth weight after autologous fresh transfer when compared with autologous FET (21). Furthermore, a retrospective cohort study compared births from fresh donor cycles and donor FET cycles in a set of recipients with at least one live birth from each transfer type and found no significant differences in birth weight or perinatal outcomes (30). Each of these findings in oocyte donation cycles contradicts the hypothesis of a significant embryonic effect of cryopreservation on birth weight.

However, retrospective studies, including registry studies, have inherent potential confounding through lack of randomization, and registry studies typically lack detail regarding treatment protocols, such as embryo culture conditions, cryopreservation technique, endometrial preparation, luteal support, and COS protocol, among other potentially relevant parameters. For example, FET cycles have frequently used supernumerary embryos after the morphologically best (primary) embryos were transferred fresh, potentially creating confounding biases in embryo selection and patient selection in retrospective and registry studies. The multiple opportunities for confounding may cloud interpretation of results.

The current study is a post hoc analysis of two concurrent randomized trials, one in normal responders (14) and the other in high responders (31), in which patients were randomly assigned to fresh ET or FET at a single center. The two studies used identical staff, culture conditions, COS protocols and medications, cryopreservation methods, embryo selection and transfer techniques, and luteal support. The inclusion criteria were similar (first-time IVF patients, 18–40 years of age, day 3 FSH <10 IU/L, no preimplantation genetic testing) and differed only in the number of antral follicles so that the normal-responder study (14) specified 8 to 15 antral follicles and the high-responder study (31) specified >15 antral follicles. The purpose of the difference in antral follicle count was so that the normal-responder protocol could specify an ovulatory trigger of hCG alone, while, for safety reasons, the high-responder study specified a trigger of low-dose hCG in combination with GnRH agonist. However, both of these trigger types were allowed in both studies, based on physician discretion. Therefore the two studies form a continuum of subjects with normal to high ovarian response treated under consistent protocols and laboratory conditions.

Both studies used only primary embryos (no supernumerary embryos) regardless of whether the transfers were fresh or frozen-thawed. The transferred embryos in the FET arms were blastocysts derived from thawed bipronuclear oocytes, so that the morphologically best blastocysts in each study arm were selected for transfer.

A post hoc analysis of these two studies will therefore be useful in examining birth weight differences, if any, without potential confounding from unknown methodological differences or nonrandomized treatment assignment.

MATERIALS AND METHODS

The two randomized trials were performed from 2007 through 2010 and included 259 subjects randomized 1:1 to cohort cryopreservation or else fresh embryo culture at a single infertility center. The two trials were registered on ClinicalTrials.gov with trial numbers NCT00963625 (normal-responder study) and NCT00963079 (high-responder study). Both studies were Institutional Review Board approved and independently monitored. Each study specified an interim stopping point and stopping criteria. The normal-responder study was stopped at that point because the stopping criteria were met (a difference in clinical pregnancy rate with P < .03), while the high-responder study was halted at that same point for safety reasons (the routine transfer of two blastocysts in all patients could no longer be supported owing to excessive twins).

Inclusion criteria specified women age 18–40 years seeking their first IVF treatment, with 8–15 antral follicles in the normal-responder study and 16 or more antral follicles in the high-responder study. Cycle day 3 FSH ≥ 10 IU/L was exclusionary, as was any use of preimplantation genetic testing. For the current post hoc study of birth weights, only live births are included in the birth weight analyses.

After informed consent, all subjects underwent conventional COS with gonadotropins, including a combination of urinary FSH (Menopur, Ferring Pharmaceuticals Inc.) and recombinant FSH (Follistim, Schering-Plough Inc.) in all cases. Follicular development was ultrasonically monitored at 2- to 3-day intervals. On or about the sixth day of stimulation, GnRH antagonist (ganirelix acetate, Schering-Plough) was initiated in all patients and sustained until the completion of stimulation.

After at least three follicles reached 18 mm in mean diameter, an ovulatory trigger of either hCG alone or else a reduced dose of hCG in combination with 4 mg GnRH agonist (leuprolide acetate) was administered to promote final oocyte maturation. The high-responder study protocol specified a dual trigger, and the normal-responder study protocol specified a trigger of hCG alone, but variation in the triggers was allowed for safety reasons, and this variation is described in the Results section. For safety reasons, patients who had extreme response to COS (typically more than 30–40 developing follicles) received GnRH agonist trigger alone and were dropped from the studies. Retrieval was scheduled 35–36 hours after trigger. In both studies, randomization was performed by
nonmedical staff drawing identical, unmarked, opaque envelopes immediately after the confirmation that at least one oocyte was collected. All collected oocytes were then inseminated by intracytoplasmic sperm injection.

Subjects randomly assigned to embryo cohort cryopreservation had all of their bipronuclear oocytes frozen by a conventional slow freezing technique described elsewhere in detail (14). In a later menstrual cycle, their entire cohorts of frozen bipronuclear oocytes were thawed at room temperature for 20 minutes in Embryo Thaw Media (Irvine Scientific).

In all cases, fresh or thawed bipronuclear oocytes were cultured to the blastocyst stage in sequential media (Quinn’s Advantage Protein Plus Cleavage Media and Quinn’s Advantage Protein Plus Blastocyst Media; Sage) before the best two blastocysts (whenever available) were morphologically selected for ultrasound-guided transfer. Patients receiving thawed ET were down-regulated with GnRH agonist (leuprolide acetate), received 10–14 days of oral E2 (Estrace, 6.0 mg daily) and E2 patches as needed to achieve an endometrial thickness ≥8 mm before P start, and initiated daily P injections (typically 100 mg) on the day before bipronuclear oocyte thaw. Patients randomized to fresh transfer initiated daily P injections 1–2 days after retrieval and E2 supplements as needed. In all subjects, P and E2 supplements were continued to sustain serum levels of at least 15 ng/mL and 200 pg/mL, respectively, until a negative pregnancy test, a pregnancy loss, or until rising serum levels demonstrated adequate placental production at approximately 10 weeks’ gestation.

Birth information, including dates and weights, was reported by patients or obtained from obstetric records. A vanished twin was identified by the number of ultrasonically observed fetal hearts exceeding the number delivered. Ultrasound findings were performed at this center.

Birth weight was the main outcome measure for this post hoc analysis. Wilcoxon’s test was used to compare numeric measures, and Fisher’s exact test was used to compare nominal outcome measures. All tests were two-tailed, and P < .05 was considered statistically significant. Linear regression models (more technically, analysis of covariance, as numeric and dichotomous nominal variables were among the available independent variables) were developed using forward selection (P < .25 to enter) followed by backward elimination (P < .05 to remain). For regression analysis, the available variables were transfer type (fresh vs. FET), maternal parameters (age at retrieval, weight, height, body mass index, tobacco use, antral follicle count), number transferred, presence of a vanished twin, and number delivered. All analyses were performed with JMP version 7 (SAS Inc.).

**RESULTS**

The two combined randomized trials included 130 subjects randomized to embryo cohort cryopreservation and 129 randomized to fresh embryo culture. Of these, 99 had FET, and 105 had fresh ET. Live birth was achieved in 74 out of 99 FETs (74.7%) and in 60 of 105 fresh transfers (57.1%), for a total of 134 live births. The live-birth rate per transfer was significantly greater with FET than with fresh transfer (P = .0119). The relative risk of failure to achieve live birth in fresh transfer compared with FET was 1.69 (95% confidence interval [CI], 1.13–2.54). The live-birth rates per randomized subject were 56.9% in subjects randomized to cohort cryopreservation and 45.5% in those randomized to fresh embryo culture (P = .1065), and the relative risk of failure to achieve live birth among patients randomized to fresh culture compared with those randomized to cohort cryopreservation was 1.24 (95% CI, 0.96–1.60).

Table 1 describes and compares cycles resulting in live birth with fresh ETs and FETs. No significant differences were observed in maternal age at retrieval, maternal weight, maternal height, maternal body mass index, tobacco use, antral follicle count, duration of stimulation, total FSH dose, serum E2 level on the day of trigger, trigger medication, number of collected oocytes, number of mature oocytes, frequency of vanished twins, number of delivered infants, or the number of singletons, twins, or triplets. There were no significant differences in subsequent pregnancy outcomes.

**Table 1**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Fresh transfer</th>
<th>FET</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Births, n</td>
<td>60</td>
<td>74</td>
<td>.8241</td>
</tr>
<tr>
<td>Maternal age, y</td>
<td>31.3 ± 3.6</td>
<td>31.3 ± 4.0</td>
<td>.6111</td>
</tr>
<tr>
<td>Maternal weight, kg</td>
<td>69.5 ± 16.3</td>
<td>70.3 ± 17.6</td>
<td>.8086</td>
</tr>
<tr>
<td>Maternal height, cm</td>
<td>165 ±8</td>
<td>164 ± 7</td>
<td>.1810</td>
</tr>
<tr>
<td>Maternal BMI, kg/m²</td>
<td>25.3 ± 5.3</td>
<td>26.1 ± 5.9</td>
<td>.5647</td>
</tr>
<tr>
<td>Maternal tobacco user</td>
<td>12/60</td>
<td>18/74</td>
<td>.6775</td>
</tr>
<tr>
<td>Antral follicles, n</td>
<td>18.8 ± 9.3</td>
<td>18.3 ± 7.3</td>
<td>.8986</td>
</tr>
<tr>
<td>Total FSH dose, IU</td>
<td>2,528 ± 745</td>
<td>2,649 ± 828</td>
<td>.3252</td>
</tr>
<tr>
<td>Stimulation period, d</td>
<td>10.2 ± 1.3</td>
<td>10.3 ± 1.2</td>
<td>.6207</td>
</tr>
<tr>
<td>E₂, day of trigger, pg/mL</td>
<td>4,081 ± 1,627</td>
<td>4,000 ± 2,277</td>
<td>.3361</td>
</tr>
<tr>
<td>Type of trigger, n</td>
<td>16/60</td>
<td>29/74</td>
<td>.1442</td>
</tr>
<tr>
<td>Dual trigger</td>
<td>44/60</td>
<td>45/74</td>
<td>.1442</td>
</tr>
<tr>
<td>Oocytes collected, n</td>
<td>17.9 ± 8.1</td>
<td>17.7 ± 8.3</td>
<td>.7555</td>
</tr>
<tr>
<td>Mature oocytes collected, n</td>
<td>14.0 ± 6.0</td>
<td>13.9 ± 6.6</td>
<td>.7132</td>
</tr>
<tr>
<td>Blastocysts transferred, n</td>
<td>1.92 ± 0.28</td>
<td>1.88 ± 0.33</td>
<td>.4729</td>
</tr>
<tr>
<td>Vanished twin, n</td>
<td>7/60</td>
<td>10/74</td>
<td>.7996</td>
</tr>
<tr>
<td>Delivered, n</td>
<td>1.42 ± 0.53</td>
<td>1.46 ± 0.53</td>
<td>.6123</td>
</tr>
<tr>
<td>Mean birth weight, kg</td>
<td>2,739 ± 0.655</td>
<td>2,873 ± 0.775</td>
<td>.2632</td>
</tr>
<tr>
<td>Gestational age at birth, d</td>
<td>258.5 ± 20.3</td>
<td>258.0 ± 22.6</td>
<td>.9750</td>
</tr>
<tr>
<td>Singleton births</td>
<td>36</td>
<td>41</td>
<td></td>
</tr>
<tr>
<td>Mean birth weight, kg</td>
<td>3.076 ± 0.511</td>
<td>3.242 ± 0.701</td>
<td>.1572</td>
</tr>
<tr>
<td>Gestational age at birth, d</td>
<td>268.3 ± 12.7</td>
<td>268.3 ± 13.4</td>
<td>.9593</td>
</tr>
<tr>
<td>Twin births</td>
<td>23</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>Mean birth weight, kg</td>
<td>2.246 ± 0.516</td>
<td>2.422 ± 0.611</td>
<td>.1143</td>
</tr>
<tr>
<td>Gestational age at birth, d</td>
<td>244.0 ± 21.3</td>
<td>245.5 ± 25.6</td>
<td>.3059</td>
</tr>
<tr>
<td>Triplet births, n</td>
<td>1</td>
<td>1</td>
<td>NA</td>
</tr>
<tr>
<td>Mean birth weight, kg</td>
<td>1.928</td>
<td>2.165</td>
<td>NA</td>
</tr>
<tr>
<td>Gestational age at birth, d</td>
<td>242.6</td>
<td>234.5</td>
<td>NA</td>
</tr>
</tbody>
</table>

significant differences in mean birth weights or gestational age at birth between fresh transfer and FET, nor when broken out by the number delivered (Table 1).

Because of the potential for multiple confounding variables affecting birth weight, most notably gestational age at birth, multiple linear regression was used to investigate in greater detail whether any of the available variables might explain the variation in mean birth weight.

After forward stepwise selection and backward elimination, multiple linear regression identified the following significant predictors of mean birth weight: gestational age at birth ($P < .0001$), number of delivered infants ($P = .0002$), vanished twin ($P = .0060$), and FET ($P = .0094$). Each additional day of gestational age was estimated to increase mean birth weight by $24.8$ g (95% CI, 21.3–28.3 g), each increment in the number delivered was estimated to reduce birth weight by $29.1$ g (95% CI, 14.2–44.0 g), a vanished twin was estimated to reduce birth weight by $282$ g (95% CI, 84–480 g), and FET increased birth weight by $166$ g (95% CI, 43–290 g) when compared with fresh transfer.

DISCUSSION
Birth weight differences were not significant in the univariate analyses in Table 1. While mean birth weights were consistently $166–237$ g greater in FET than in fresh transfer across each number delivered, none of these comparisons were statistically significant. A multivariate procedure, multiple linear regression, was employed to discern the contributions of each significant variable.

This is the first report comparing birth weights in patients randomly assigned to fresh or FET. Many prior studies have investigated the potential effects of COS and cryopreservation on birth weight. These have all been retrospective studies, either of national registries or of results at individual centers. Studies of national registries have very large sample sizes but lack detail regarding treatment assignment and medical methodology, and retrospective studies generally confer substantial opportunities for confounding effects through lack of randomization. For example, a common weakness in retrospective studies comparing fresh transfer and FET is the potential that the former includes only primary embryos, while the latter is mainly supernumerary embryos after the morphologically best embryos were transferred fresh.

The current finding suggests there were no substantial confounding uncontrolled variables causing the birth weight differences in the prior studies. The current study, although small in sample size, has the advantages of known and detailed methodology and randomized treatment assignment. For example, all patients in both groups had the same embryo culture conditions and had only their primary (morphologically best) embryos selected for transfer, and none had genetic screening of embryos.

The results were consistent with the conclusions of many prior retrospective studies (16–29), suggesting those prior findings were not due to confounded uncontrolled variables. Specifically, the current study found $166$ g (95% CI) greater birth weight after FET when compared with fresh transfer, a value very near the center of the range of $80–250$ g difference reported in 10 registry studies (16–25) and very close to the $156$ g reported value from analysis of a U.S. registry (21).

The two randomized trials combined here had very similar inclusion criteria and methodology. The only methodological difference in the protocols was that the high-responder study specified a dual trigger as the primary therapy to reduce ovarian hyperstimulation syndrome risk, while the normal-responder study specified hCG alone. However, both trigger types were allowed in both studies per physician judgment, and the normal-responder study also included 20 births that followed dual triggers after a greater than expected ovarian response, while the high-responder study included one live birth that followed an hCG trigger in a patient with lower ovarian response than expected. Therefore, there was no clear distinction in methodology between the two trials, and their combined data set is actually a continuum of subjects with at least eight antral follicles. The proportion of use of each trigger did not differ significantly between the fresh and FET groups, and the randomization always occurred immediately after oocyte retrieval so that the physician’s trigger choice could not be influenced by treatment assignment.

The current study cannot resolve whether the observed birth weight difference resulted from an embryonic effect of cryopreservation or an endometrial effect of COS. However, comparisons of birth weights in cycles using embryos derived from donor oocytes have reported no difference in birth weight or in the incidence of low birth weight between fresh transfers and FET, contradicting the hypothesis of an embryonic effect of cryopreservation on birth weight in the absence of uterine COS exposure (21, 30).

The cryopreservation protocol used here was conventional slow freezing of bipronuclear oocytes, followed by thaw, extended culture, and blastocyst transfer. This method is associated with reduced blastocyst yield when compared with extended culture from fresh oocytes (32) and has been largely replaced by blastocyst vitrification (33). However, the resulting blastocysts from this technique have implantation potential comparable to that of vitrified-thawed blastocysts, and the artificial endometrial preparation is similar in both techniques (33). Should there prove to be any embryonic effects of cryopreservation on birth weight, then the results here might be method specific and not specifically relevant to blastocyst vitrification. However, as discussed above, current evidence suggests the frequently observed birth weight differences between fresh and thawed ETs result from uterine effects, and therefore the results presented here and in the cited registry studies should be relevant to other cryopreservation protocols.

The live-birth rate per transfer was significantly greater with FET than with fresh transfer, suggesting inferior endometrial receptivity in fresh transfer after COS exposure as concluded before (14). The causal mechanisms behind the altered endometrial receptivity may be related to those affecting birth weight. However, the live-birth rate per randomized subject did not differ significantly. Therefore more evidence is needed to determine whether cohort
cryopreservation improves the chance of live birth. However, the birth weight differences reported here and in the many cited registry studies remain a significant factor to consider.

REFERENCES


33. Shapiro BS, Daneshmand ST, Garner FC, Aguirre M, Hudson C. Freeze-all at the blastocyst or bipronuclear stage: a randomized clinical trial. Fertil Steril 2015;104:1138–44.