What is the most relevant standard of success in assisted reproduction?

No single outcome measure is satisfactory when evaluating success in assisted reproduction; both twin births and singleton births should be counted as successes

Richard P. Dickey¹ ²*, Belinda M. Sartor¹ and Roman Pyrzak¹

¹The Fertility Institute of New Orleans, New Orleans, Louisiana, ²Section of Reproductive Endocrinology, Department of Obstetrics and Gynecology, Louisiana State University School of Medicine, New Orleans, Louisiana.*Corresponding author at: The Fertility Institute of New Orleans, 6020 Bullard Avenue, New Orleans, Louisiana 70128. Tel.: 504 246 8971, Fax: 504 246 9778, E-mail: info@fertilityinstitute.com

The 2002 recommendation of the consensus meeting of the European Society of Human Reproduction and Embryology (ESHRE) that the outcome measure of assisted reproductive technology (ART) and non-ART should be ‘singleton live birth rate’ could profoundly effect the ability of infertility patients to become pregnant. We reviewed published reports and new data concerning elective single embryo transfer (eSET) vs. double embryo transfer (DET) and the outcome of twin pregnancies in the United States, as well as recommendations of other Societies concerning number of embryos to transfer and methods used to measure ART success. We found that no single outcome measure of ART is ideal. Mandatory eSET would result in 42%±70% fewer births compared to DET. Infertility treatments account for only 12% of all twin pregnancies and 4% of all premature births in the United States. Twin and singleton births due to ART do not occur earlier than spontaneously conceived twins and singletons unless they started as triplet and higher order pregnancies. Multiple outcome measures are necessary when evaluating ART success. Twin as well as singleton births should be counted as ART successes. The essential aim of infertility treatment should be a healthy low order (singleton or twin) birth.

Key words: assisted reproductive technology; in vitro fertilization; outcome measures; twin births; singleton births

Introduction

The conclusion of the consensus meeting of the European Society of Human Reproduction and Embryology (ESHRE) in May 2002 that “…the essential aim of IVF/ICSI is the birth of one single healthy child, with a twin pregnancy being regarded as a complication”, made inevitable the meeting’s recommendation that ‘the outcome measure of assisted reproductive technology (ART) and non-ART should be ‘singleton live birth rate’ (Land and Evers, 2003). This conclusion, and the recommendation that only single live births should be counted when measuring the success of infertility treatment, must be evaluated in terms of their effect on individual infertility couples and on utilization of national health resources.

The decision that a twin birth is a complication and that the only acceptable outcome of infertility treatment is a single live birth, is unnecessary, unsympathetic to couples who require ART or ovulation induction (OI) in order to achieve pregnancy, and based on incomplete information. In order to achieve this result, by which the quality of their program is measured, ART clinics would need to perform elective single embryo transfer (eSET) in all IVF cycles, and OI with clomiphene citrate (CC) or urinary and recombinant gonadotropins (GT), would have to be abandoned entirely. A singleton birth policy for ART and non-ART will multiply cost and discomfort for couples who require IVF, desire two children, and have no physical impediment to successful completion of a twin pregnancy. Extension of this policy to OI outside of ART will result in denial of pregnancy to many infertile couples who cannot conceive without OI and who are unable to have IVF because of economic or other reasons.

Single versus double embryo transfer

Infertility treatment contribution to twin and higher order multiple births

Multiple pregnancies may represent the dark side of modern infertility treatment, but the problems associated with twin births, which are natural phenomena, are minimal compared to the complications due to triplet and higher order pregnancies. Moreover, infertility treatment is responsible for only a small part of the total twin births that occur annually. In 2000, IVF and other ART procedures were responsible for 12% of 118 997 babies born as twins compared to 42% of 7328 triplet and higher order births in the US (Reynolds et al., 2003). During the same year, OI without ART was estimated to be responsible for 21% of twin births and 40% of triplet and higher order
births, whereas 67% of twin births and 18% of triplet and higher order births were estimated to result from natural conception (Reynolds et al., 2003). Limiting the number of embryos transferred by eSET would reduce triplet and higher order births from all causes by 42%, but would reduce total twin births by no more than 12%, and some twin births due to IVF would still occur because of monozygotic splitting. Limiting the number of embryos transferred to two by elective double embryo transfer (eDET) would also eliminate 42% of triplet births, and the number of twin births due to IVF would be decreased by ~20%, by eliminating twin births that begin as triplet and higher order gestations. Clearly, eDET would reduce the number of high order multiple births, which have the greatest risk of complications, as effectively as eSET.

Triplet and higher order pregnancies are one reason ART singletons and twins are born prematurely

A retrospective analysis of multiple pregnancies, diagnosed by ultrasound before the 7th week of gestation, found that 47% of triplet gestations and 35% of quadruplet gestations had undergone spontaneous reduction or had lost fetal heart activity by the 12th week of gestation, and were continuing as viable singletons or twins (Dickey et al., 2002). Additional unpublished analysis, presented in Table I, shows that 15% of singleton births and 19% of twin births following IVF began as higher order gestations. Singleton births following IVF that began as single, twin, and triplet or higher order gestations were born 1.8 days, 5.0 days, and 12.4 days earlier, respectively, compared to spontaneous singleton births that began as single gestations. Twin births, due to IVF, that began as twin, triplet, and quadruplet or higher order gestations were born 2.7 days, 6.2 days, and 18 days earlier, respectively, compared to spontaneous twin births that began as twin gestations. In the absence of spontaneous reduction from higher order gestations, the proportion of singletons due to IVF, born before 32 weeks, was not increased compared to spontaneously conceived singletons (1.1 versus 1.3%). Similarly the proportion of premature twins was only slightly increased (10.4 versus 9.5%) compared with spontaneously conceived twins.

This information, which has not been presented previously, indicates that the increased incidence of premature birth reported for IVF singleton and twin births, compared to spontaneous pregnancies, is due in large part to the initial occurrence of triplet and higher order gestations. Before accepting that twin births due to IVF are at inherently greater risk for prematurity than spontaneous twin births, on the basis of data from IVF procedures performed before 1998, when transfer of three or more embryos was common, it is necessary to reevaluate the outcome of IVF births that followed transfer of two embryos, into patients of similar age and parity.

Twins due to ART/ICSI and OI account for a small proportion of all premature births

The increased incidence of perinatal mortality and later developmental abnormalities of twins, compared to singleton births, are principally related to a higher incidence of prematurity. However, twins whether natural or the result of infertility treatment, represent only 12% of all premature births in the US (Gardner et al., 1995). Twins due to ART, therefore, account for only 1.4% of total premature births in the US, and twins due to OI outside of ART account for only an additional 2.5% of total premature births. Furthermore, infants from multiple births have a greater chance of survival than singleton infants, of the same birth weight, gestational age, and ethnic origin (Draper et al., 1999).

In the United States, babies born at 34 weeks develop normally and have few neonatal problems. Whereas twins as a group are at high risk for preterm birth, some women are clearly able to carry twins to 34 weeks and beyond without difficulty. We suggest that rather than requiring eSET for all ART patients and forbidding OI outside of ART because use of

<table>
<thead>
<tr>
<th>No. sacs</th>
<th>Births/treatment</th>
<th>Length gestation, days</th>
<th>Pre-term, 24–32 weeks</th>
<th>&lt;37 weeks</th>
<th>≥37 weeks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single Birth</td>
<td>None</td>
<td>2680</td>
<td>276.6</td>
<td>–3.4</td>
<td>34</td>
</tr>
<tr>
<td>Ivy/ART</td>
<td>1 GS</td>
<td>261</td>
<td>274.8</td>
<td>–5.2</td>
<td>3</td>
</tr>
<tr>
<td>2 GS</td>
<td>41</td>
<td>271.6</td>
<td>–8.4</td>
<td>1</td>
<td>2.4</td>
</tr>
<tr>
<td>≥3 GS</td>
<td>5</td>
<td>264.1</td>
<td>–15.8</td>
<td>0</td>
<td>–</td>
</tr>
<tr>
<td>Twin Birth</td>
<td>None</td>
<td>21</td>
<td>254.5</td>
<td>–25.5</td>
<td>2</td>
</tr>
<tr>
<td>Ivy/ART</td>
<td>2 GS</td>
<td>106</td>
<td>251.8</td>
<td>–28.2</td>
<td>11</td>
</tr>
<tr>
<td>3 GS</td>
<td>21</td>
<td>248.3</td>
<td>–31.7</td>
<td>3</td>
<td>14.3</td>
</tr>
<tr>
<td>≥4 GS</td>
<td>4</td>
<td>236.5</td>
<td>–43.5</td>
<td>1</td>
<td>25.0</td>
</tr>
</tbody>
</table>

1 From adjusted onset of menstruation; assumes ovulation, or oocytes retrieval day 14.
2 <37 weeks = 24.0 to 36 weeks inclusive.
CC and GT might result in multiple pregnancies, it may be possible to identify women who would be at increased risk if they delivered twins so that IVF with eSET can be preferentially provided for those patients. Factors known to be associated with premature birth in multiple pregnancy include nulliparity (Blickstein et al., 2000), short stature (Blickstein et al., 2003), and cervical length at 24–26 weeks gestation (Imseis et al., 1997).

**ESHRE, WHO and ASRM recommendations**

It is an encouraging that as yet, neither ESHRE, nor the World Health Organization (WHO) requires eSET for all patients. ESHRE recommendations state, ‘eSET should be performed if a twin pregnancy is contraindicated and/or if a couple wishes to avoid a twin pregnancy at any cost’, and ‘eSET should be proposed in a first or second IVF/ICSI cycle in women <36 years of age if at least one good quality embryo is available’ (Land and Evers, 2003). The WHO recommended in September 2001 that ‘no more than two embryos be transferred per cycle’ (Vayena et al., 2001). The American Society for Reproductive Medicine (ASRM) guidelines recommend: ‘transferring a maximum of two embryos for age <35, when there are at least three ‘improved’ quality embryos and excess embryos available for cryopreservation; transferring a maximum of three embryos for age <35, when there are no more than three good quality embryos with none available for cryopreservation; transferring a maximum of four embryos for age 35 to 40; and transferring a maximum of five embryos for age 40 and above, or for multiple failed cycles’ (American Society for Reproductive Medicine, 1999).

**eSET results in significantly lower birth rates than eDET**

An ESHRE Campus Course Report (2001), summarizing published data on eSET showed per cycle clinical pregnancy rates for eSET equal to rates for eDET. In contrast to those findings, the United States national IVF clinic results for 2001, provides unequivocal evidence that birth rates for eSET are significantly lower than for eDET, and also that birth rates are not increased when more that two fresh non-donor embryos are transferred (Centers for Disease Control and Prevention et al., 2003). The live birth rate per transfer, for patients with the most favorable prognosis, because of age less than 35 years with extra embryos cryopreserved for future use, was 30.0% for eSET and 51.7% for eDET. The live birth rate for all 387 clinics and 65,363 transfer cycles, performed during 2001 in the United States, was 11.3% for SET and 37.2% for DET. When three, four, or more than four embryos were transferred the live birth rate was less than 47% for patients with the most favorable prognosis, and less than 35% for all patients. These results indicate that, if eSET had been mandatory, the live birth rate for patients with the most favorable prognosis would have been reduced by 42%, and the birth rate for all patients might have been reduced by 70%. Until results from other national and multi-national databases, that include all patients as well as patients with favorable prognosis, are available, this data must be considered to represent the real consequences if a single embryo transfer policy was enforced for all patients.

How should we measure ART success?

The denominator for measuring ART success cannot be agreed on

ESHRE’s recommendation that ‘the outcome measure of assisted reproductive technecology (ART) and non-ART should be a ‘singleton live birth rate’ (Land and Evers, 2003), does not specify a denominator. Whether this all-important statistic should be per cycle, per retrieval, per transfer, or per embryo transferred is left unstated. Others have supplied this deficiency according to their own objectives. Min et al. (2004) champion the singleton, term gestation, live birth rate per cycle begun as the measure of ART success. In an earlier issue of this Journal, Vail and Gardner (2003) stated that live birth rates per person are the only appropriate way of reporting the results of IVF and other infertility treatment. The WHO has recommended that ‘The principal outcome statistic of ART results should be singleton and multiple live-birth rates per treatment cycle initiated’ (our emphasis) (Vayena et al., 2001).

**National ART reporting systems**

National reporting systems are designed not only to gather health statistics and monitor clinical activity but also to allow couples to determine their own possibility of pregnancy and which ART programs offer them the best chance for success. In Sweden, two independent surveillance systems are employed, the one concerned with clinical specific results, the other concerned with follow-up of the health of IVF children (Nygren 2002). Swedish clinical results include the proportion of singleton, duplex and triplet or higher order deliveries per cycle, per ovum aspiration, and per embryo transfer in five-year age intervals, the number of embryos replaced in relation to pregnancy outcome, and the proportion of children born with low birth weight. In the United States, the national reporting system is similar to Sweden’s except that the percent of children born with low birth weight, and the clinical specific pregnancy rate per embryo transferred is not reported, and age intervals are different (Schieve et al., 2002). In the United States, clinic-specific information is available to everyone on the internet at http://www.cdc.gov/nccdphp/drh/art.htm

**Comparisons of individual ART programs based on births rates per cycle are subject to bias**

In the United States and elsewhere, some ART programs withhold IVF treatment from patients who they believe have a poor prognosis because of age ≥38 years, low antral follicle counts, or elevated basal FSH levels, and urge oocyte donation or adoption instead (Bansci et al., 2002). Other programs provide treatment to similar patients even though they are more likely to have cycles cancelled before oocyte retrieval, and are less likely to become pregnant if they do have oocyte retrieval (Fratterelli et al., 2000; Bansci et al., 2002). ART programs differ concerning cancellation policies when fewer than four preovulatory follicles develop; some programs proceed to retrieval when there are as few as one or two follicles. Most ART programs in the United States offer a trial of OI with CC or GT and intrauterine insemination (IUI) to patients without

785
tubal obstruction, but other programs take similar patients directly to IVF. The percentage of second, third, and latter ART cycles differs between programs. Success rates per cycle initiated, or per intent to treat, cannot reliably indicate the comparative success of ART programs nor the probability that an individual patient will become pregnant, unless only first ART cycles are compared, both the number of antral follicles and age are identical, patients have received the same treatment previously, and cancellation and transfer policies are similar.

No single measure of ART success is ideal, therefore outcome reports should include as much information as possible

Measurements of ART success that are independent of patient characteristics and patient selection or cancellation policies, are required to compare ART programs. The effect of antral follicle numbers and pretreatment screening, but not previous treatment, can be mitigated to some extent by measuring success as live birth rate per retrieval or per embryo transfer. Other measures that have been proposed are the number of oocytes retrieved (Fratterelli et al., 2000), number of quality embryos per oocyte, and implantation rate. Because none of these are ideal, more than one measure must be used.

Outcome measures should include all healthy live births, both single and multiple, per cycle with the percentage of premature births indicated for each. The percentage of cycle cancellations and number of embryos cryopreserved should be reported, because these indicate the proportion of patients with high and low prognosis for success admitted to treatment. Outcome measures used to compare ART programs must at a minimum be stratified by age and antral follicle number, number of previous ART and non ART cycles, and should exclude patients with excessive FSH levels. The number of mature follicles per cycle initiated, stratified by patient age and antral follicle count, in patients who have not had previous cycles of OI or ART, is possibly the most accurate outcome measure with which to compare drug stimulation regimens. The live birth rate, both twin and singleton, at 34 weeks and later, per first oocyte retrieval cycle, stratified by age, antral follicle count, semen quality, tubal status, and previous treatment, is arguably the most accurate outcome measure with which to compare individual ART programs.

In the future, preimplantation genetic diagnosis (PGD) may make single embryo transfer practical for all patients because it offers the possibility of identifying and transferring only embryos that are chromosomally normal. For the present, only the most serious genetic defects are being diagnosed by PGD and the procedure is as yet too expensive to allow use in all IVF/ICSI (Robertson, 2003). If PGD becomes widely used, it will be necessary to report outcome separately for patients who have PGD before transfer.

Conclusions

In conclusion, we contend that the essential aim of IVF/ICSI should be a healthy low order (singleton or twin) birth, and that there is no one best outcome standard for success in ART. High order multiple births due to ART, which concern all reasonable persons, can be avoided by strictly limiting the number of embryos transferred to a maximum of two embryos. Twin pregnancy should be regarded as a necessary but manageable complication of infertility treatment for patients who require ART or OI. Patients undergoing ART should have the option of double embryo transfer providing they are healthy, able to carry a twin pregnancy to 34 weeks and desire more than one child. ART programs should not be penalized for providing patients the option of double embryo transfer, by not counting twin births when reporting IVF ‘success’.

References


Submitted: 11 November, 2003; accepted: 16 December 2003