Redefining in vitro fertilization success: should triplets be considered failures?

Lawrence Grunfeld, M.D.,[a,b] *Martha Luna, M.D.,*[a,b] *Tanmoy Mukherjee, M.D.,*[a,b] *Benjamin Sandler, M.D.,*[a,b] *Yui Nagashima, B.A.,*[b] *and Alan B. Copperman, M.D.*[a,b]

[a] Department of Obstetrics and Gynecology, and Department of Reproductive Endocrinology and Infertility, Mount Sinai School of Medicine; and [b] Reproductive Medicine Associates of New York, New York, New York

Objective: To evaluate how the ranking of IVF programs changes if high-order multiple pregnancies (HOMPs) are considered negative outcomes.

Design: Retrospective analysis.

Setting: The 2004 Society for Assisted Reproductive Technology Clinic Outcome Reporting System.

Patient(s): Two hundred seven programs that performed >50 IVF cycles were analyzed, and only patients ≤35 years of age were included.

Intervention(s): Program size, success rate, and number of embryos transferred were recorded for each program. All programs were ranked in accordance to these variables.

Main Outcome Measure(s): The programs were re-ranked after considering HOMPs as negative, rather than as positive, outcomes.

Result(s): High-order multiple pregnancies were more abundant as the number of embryos replaced increased. The live birth rate, after excluding HOMPs, significantly decreased as the number of embryos transferred increased. The lowest ranked programs performed transfers with the greatest number of embryos (2.4–4.5). After HOMP exclusion, the rank of programs that transferred >2.4 embryos decreased, as opposed to the programs that transferred fewer embryos (1.8–2.4).

Conclusion(s): Reclassifying HOMPs as failed cycles will result in a reduced rank in programs that maintain high embryo replacements. Redefining success as a singleton or twin birth significantly changes the ranking order of programs and, potentially, how programs are perceived. (Fertil Steril® 2008;90:1064–8. ©2008 by American Society for Reproductive Medicine.)

Key Words: IVF success, triplets, high order multiple pregnancies, live birth rate, multiple gestation

The term *IVF success rates* provokes strong emotions on the part of reproductive specialists. Despite the admonition of the Society for Assisted Reproductive Technologies (SART) that "a comparison of clinic success rates may not be meaningful because patient medical characteristics and treatment approaches may vary from clinic to clinic," it is inevitable that patients and physicians will use the registry for direct comparisons. In general, the majority of programs use the SART guideline that a successful IVF outcome is a "live birth per IVF procedure." Programs may, in an attempt to improve apparent success rates, transfer a greater number of embryos. There is, however, a plateau effect that varies from program to program, in that greater numbers of embryos will increase the multiple-pregnancy rates without a major improvement in success rates (1–4). The actual physical, emotional, and financial costs of this strategy may be far greater than realized because of the increased morbidity that is associated with multiple pregnancies (5–7).

Interwoven with the issue of competing success rates is the multiple-gestation rate that currently is reaching epidemic proportions in the United States. Although other countries have effectively reduced multiple gestations by legislative changes, the United States still is dependent upon voluntary adherence to the American Society for Reproductive Medicine guidelines. The problem is now so significant that 35% of deliveries after assisted reproductive technology were multiple births, as compared with 3% in the general United States (8), and the 2004 SART Clinic Outcome Reporting System online report, which documented a 4.9% triplet rate in women who were <35 years of age prior to using their own oocytes. Given that multiple gestation is the greatest risk factor for premature delivery, it is logical to conclude that a cavalier attitude toward multiples will lead to significant future neonatal morbidity, if not mortality.

The current method of calculating IVF success rates considers multiple gestations and prematurely delivered infants to be successes, rather than failures. Patients may see multiple births as an acceptable outcome and may even construe multiple gestations as actualizing a better return on their financial investment. They may be less sophisticated about the greater risk of neonatal morbidity and pregnancy loss. When patients evaluate IVF programs on the basis of Centers for Disease Control and Prevention–reported data, they may conclude that a program with a higher success rate may be preferable to a program with a reduced success rate, even if the latter program has a very low multiple-pregnancy rate.

Received January 17, 2007; revised and accepted July 23, 2007.

Reprint requests: Alan B. Copperman, M.D., Department of Obstetrics and Gynecology, Division of Reproductive Medicine, Mount Sinai School of Medicine, 635 Madison Avenue, 10th Floor, New York, New York 10022 (FAX: 212-756-5770; E-mail: acopperman@rmany.com).

0015-0282/08/$34.00
doi:10.1016/j.fertnstert.2007.07.1360

For the purposes of the study, we chose to evaluate how reclassifying IVF success would change the current way we perceive programs. By incorporating higher order multiple-gestation rates into the calculation of success rates, we attempted to address whether reclassification would afford patients a more accurate measure of their reproductive outcome after IVF. Specifically, we asked whether reclassifying triplet pregnancies as failed cycles would result in a reduced rank in programs that maintain high embryo replacements and high multiple-pregnancy rates.

MATERIALS AND METHODS

Institutional review board approval was obtained for retrospective data collection and analysis of laboratory and clinical data. The 2004 SART Clinic Outcome Reporting System online report was used to assess the reproductive performance of the 385 assisted reproductive technology centers that report to SART. The data set is fully deidentified, in accordance with Health Insurance Portability and Accountability Act regulations, and was analyzed in the current study. Two hundred seven programs reported ≥50 cycles to SART and were used for the analysis. Smaller programs were not evaluated because they had few pregnancies. Only data for patients <35 years of age were examined. Success, for the study, was defined as live birth per embryo transfer procedure. A live birth is a cycle that resulted in at least one live-born neonate, regardless of the number of other neonates and of whether they were live born or stillborn.

The size of the programs (number of cycles performed), the success rates of the programs, and the mean number of embryos transferred were recorded, based on the traditional measure of live birth per embryo transfer. All programs were ranked in accordance to these factors and were organized, in an ascending mode, into quartiles and analyzed independently. The success rates were then recalculated, subtracting pregnancies that were diagnosed as triplet or higher-order multiple pregnancies (HOMPs), and the centers were reranked.

Statistical analysis was performed by using Analyse IT software for Microsoft Excel 2000 (Analyse-it Software, Ltd., Leeds, West Yorkshire, United Kingdom). Logistical regression, one-way analysis of variance, and χ^2 tests were used as indicated. Significance was set at the $P<.05$ level.

RESULTS

Quartiles were established on the basis of program size. Quartile 1 included 52 programs that performed between 50 and 72 cycles annually. Quartile 2 included 52 programs that performed between 73 and 109 cycles. Quartile 3 included 52 programs in which 110–180 cycles were performed annually, and finally, quartile 4 consisted of 51 large programs that performed between 181 and 1,095 cycles. According to the initial success rate, 52 programs had a success rate of 16%–36%, 52 programs had a 36%–42% success rate, 52 had a 42%–48% success rate, and the final group had a success rate of 48%–67%. On the basis of the number of embryos transferred, the programs within the first group had a mean number of 1.8–2.2, the second group had a mean of 2.2–2.4, the programs within the third group had a mean of 2.4–2.8, and the last 51 programs had a mean number of 2.8–4.5 embryos transferred.

Of the 207 programs analyzed, success rates ranged from 16.7% to 67.3% live births per embryo transfer, and triplet rates ranged from 0 (in 43 programs) to 40.7%. Fifty programs (25.1%) had a >10% triplet rate, and 15 programs (7.2%) had a >20% triplet rate.

Size of Program

The live birth rate according to the size of the program was not significantly different among the four groups. The live birth rate reported in the four groups was between 32.4% and 37.1%. After correcting the data by removing the triplet and higher-order pregnancies, the success rate decreased to 28.8%–33.8% (not statistically significant). The overall success rates, after correcting the data, dropped by 8.3% and 12.1% (not statistically significant). Success rates were not affected by the size of the programs. The triplet rate across all groups was not different according to size of the program, with an overall mean between 4.1% and 6.3% ($P=.114$; Table 1).

Success Rates

The live birth rates of programs with lower success rates did not differ from those of programs with higher success rates, after excluding the triplet and higher-order pregnancies ($P=.18$; Table 1).

Number of Embryos Transferred

We further analyzed the difference in success rates after excluding triplet pregnancies according to the number of transferred embryos. We encountered that programs that transferred the fewest embryos presented the highest success rates, both before and after data adjustment. The success rate, after excluding HOMPs, decreased significantly across all groups as the number of embryos transferred increased ($P<.001$). There was a 21.8% decrease in live birth rates when triplet pregnancies were eliminated in programs presenting the highest mean number of transferred embryos. The triplet rate was higher as the mean number of embryos replaced increased (Table 1 and Fig. 1).

The lowest ranked programs performed transfers with the greatest mean number of embryos (2.4–4.5). After excluding triplets, the mean rank of programs that transferred >2.4 embryos worsened, as opposed to the case of the programs that transferred fewer embryos (1.8–2.4), which presented an improved ranking position in terms of success (Table 2).

Thirty-two percent of the programs in which the fewest number of embryos were transferred improved their ranking position, whereas 1% of these programs declined in their

TABLE 1

Programs' size, success rates, and mean number of transferred embryos (ET), based on live birth rate (LBR) per ET.

Parameter	LBR including HOMP	LBR excluding HOMP	Mean rate of HOMP	% Decrease in LBR after data correction
Program size				
Quartile 1 (n = 52)	32.4	28.8	5.7	11.2
Quartile 2 (n = 52)	34.7	30.5	6.3	12.1
Quartile 3 (n = 52)	37.1	33.8	4.6	9
Quartile 4 (n = 51)	33.7	30.9	4.1	8.3
Success rate				
16%–36% (n = 52)	29.5	25.7	4.1	12.9
36%–42% (n = 52)	39.3	33.1	6.0	15.8
42%–48% (n = 52)	45.7	40.1	4.9	12.3
48%–67% (n = 51)	54.9	49.1	4.9	10.6
Mean no. of ETs				
1.8–2.2 (n = 52)	46.3	44.2	2.1	4.5[a]
2.2–2.4 (n = 52)	44.3	39.9	4.4	10.0[a]
2.4–2.8 (n = 52)	40.7	34.6	6.1	15.0[a]
2.8–4.5 (n = 51)	37.7	29.5	8.3	21.8[a]

Note: All data are percentages. HOMP = high-order multiple pregnancies (≥3 fetuses).
[a] $P < .001$.

Grunfeld. Redefining in vitro fertilization success. Fertil Steril 2008.

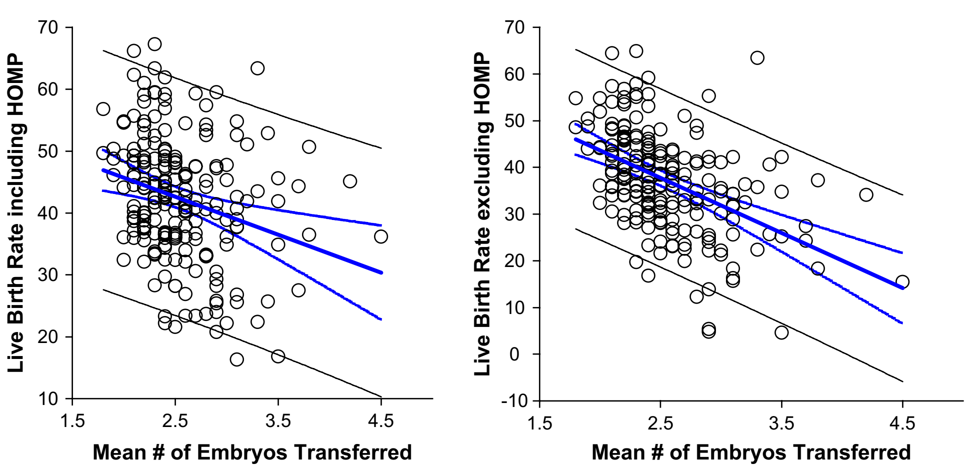
position. Eleven percent of programs that had a mean of 2.2–2.4 embryos transferred presented improvement in their rank, and 5% had a decline. The third group (a mean of 2.4–2.8 embryos transferred) presented a decrease in rank in 26% of the programs, whereas 15% of the programs improved their ranking position. Thirty-two percent of programs that had the greatest mean number of transferred embryos presented a decline in rank, and 9% of these programs improved in rank.

DISCUSSION

Since IVF was introduced, there has been a trend toward increased success, with a concomitant increase in multiple-

FIGURE 1

Programs that transferred the fewest embryos presented the highest success rates, both before and after data correction.



Grunfeld. Redefining in vitro fertilization success. Fertil Steril 2008.

TABLE 2

Numbers of programs per quartile, their mean numbers of transferred embryos (ET), their ranks, and their changes in rank after excluding HOMP.

Quartile no.	No. of programs	Mean no. of ET	Overall mean rank	New rank	Change in ranking position
1	52	1.8–2.2	80.6	62.9	+17.7
2	52	2.2–2.4	96.2	92.2	+4
3	52	2.4–2.8	115.2	121.8	−6.6
4	51	2.8–4.5	129.7	146.4	−16.7

Grunfeld. Redefining in vitro fertilization success. Fertil Steril 2008.

gestation rate. This trend has been attributable to improved stimulation regimens, improved laboratory quality, and the introduction of multiple-embryo transfers. The most frequent measure of IVF success per center continues to be number of live births per cycle, irrespective of neonatal outcome. Because HOMPs are reported to SART before discharge and fetal reduction, the incidence of these pregnancies is consistent. We have demonstrated in this study that reclassifying higher order multiple gestations as failed cycles would result in a reduced rank for programs that maintain high embryo replacements and high multiple-pregnancy rates. By redefining success as a singleton or twin birth, the ranking order of programs and potentially how individual programs are perceived would change, and patients would potentially have a better marker of identifying their chance of achieving true success at each individual IVF center.

Reproductive specialists are aware of the risks and complications related to multiple pregnancies, especially triplet and higher-order pregnancies. Multiple strategies, such as blastocyst culture, preimplantation genetic screening, embryo cryopreservation, and more recently, proteomics and metabolomics of embryonic culture media, are aiding in the goal of achieving single-embryo transfer without sacrificing success rates (9). However, because these strategies are not available to all centers and not all patients are candidates for their applicability, we face a realistic problem that is not exclusive to certain programs, countries, or continents.

According to the 2004 Centers for Disease Control and Prevention report (10), in the year 2002 in the United States, the mean number of embryos transferred was 2.7 for women <35 years of age, 3.0 for women 35–37 years of age, 3.3 for women 38–40 years of age, and 3.5 for women 41–42 years of age. Approximately 62% of fresh non-donor IVF transfers included three or more embryos transferred. The incidence of multiple-fetus pregnancy for non-donor IVF patients was 36% (29.4% twins and 6.8% triplets). Given these relatively elevated HOMP rates, when compared with spontaneous events, and given the comorbidity and mortality associated to this phenomenon, we set out to determine whether a difference in rates is encountered by eliminating these triplet pregnancies from the programs' success, based on program size, and according to the number of transferred embryos. We wanted to determine whether the programs' respective ranking positions changed by eliminating these cases.

For one, we found that overall success rates are not correlated with the size of the program, because we found that smaller size programs had results comparable to those of larger programs. Triplet rates were not different on a program-size basis, either. We did not demonstrate a correlation between triplets and the success rates of a program. However, we encountered that triplet and higher-order pregnancies are more abundant as the mean number of embryos replaced increases. In a related finding, programs that transfer higher numbers of embryos tend to have lower success rates.

In addition, when we reclassified the ranking position of these programs that have a higher mean number of transferred embryos, their ranking positions dropped significantly (17 positions). On the contrary, for those programs that transfer fewer amounts of embryos, we encountered a significant improvement in their ranking position (18 places), after eliminating triplet and higher-order pregnancies. Reclassifying triplets as failed cycles will result in a reduced rank for programs that maintain high embryo replacements and high multiple-pregnancy rates. We have demonstrated that redefining success as a singleton or twin birth significantly changes the ranking order of programs and potentially how programs are perceived. It may be wise to exclude triplet pregnancies from success when evaluating IVF center performance. This will remove the incentive to transfer higher embryo numbers in an attempt to overcome uncertainty about implantation rates. Our data demonstrate that this problem is most prevalent in programs that have poor success rates. It is possible that if this reclassification system is adopted, the consumer market will limit the number of embryos transferred, and this could potentially result in avoiding the need for legislative oversight. Concern about the higher incidence of adverse outcomes associated with multiple pregnancy has led SART to establish new guidelines recommending the number of embryos or oocytes to be transferred in certain patient groups. These guidelines are based on analysis of the US experience, as reported to SART. The impact of implementation of these guidelines, which were released in 2006, will not be known for several years (11).

In the absence of federal regulations on the number of embryos that may be transferred, it is the free market that polices IVF program policies. Increased collection of accurate and useful data will ultimately assist each individual program, as well as the consumer, in identifying practice patterns and outcome probabilities.

Many infertile couples are unaware of the risks of multiple births and hence welcome these as an outcome to their fertility treatment (12). This desire, combined with financial pressures on couples of having to fund their own IVF treatment and with market competition between fertility clinics, increases the likelihood that multiple embryos will be replaced to maximize the so-called success rate. However, that rate must also be balanced with consideration of effects on maternal and fetal health. Increasing patient awareness of the risks of multiple births through counseling is an important strategy for reducing the incidence of multiple gestations.

Performance of assisted reproductive technology centers should be measured and reported so as to place greater emphasis on the outcome of singleton or even twin pregnancies, rather than on the traditional measure of live births (regardless of the quantity) per assisted reproductive technology cycle initiated. Instead, higher order multiple gestations should be regarded as complications. Infertility specialists must play an important role in ensuring that the risk of multiple pregnancies is minimized by education of patients and health care professionals and by promoting structural change in practice. The aim of IVF treatment should shift from maximizing the pregnancy rate per treatment cycle to optimizing the number of healthy deliveries achieved per patient treated.

REFERENCES

1. Templeton A, Morris JK. Reducing the risk of multiple births by transfer of two embryos after in vitro fertilization. N Engl J Med 1998;339:573–7.
2. Roest J, van Heusden AM, Verhoeff A, Mous HV, Zeilmaker GH. A triplet pregnancy after in vitro fertilization is a procedure-related complication that should be prevented by replacement of two embryos only. Fertil Steril 1997;67:290–5.
3. Staessen C, Janssenswillen C, Van den Abbeel E, Devroey P, Van Steirteghem AC. Avoidance of triplet pregnancies by elective transfer of two good quality embryos. Hum Reprod 1993;8:1650–3.
4. Tasdemir M, Tasdemir I, Kodama H, Fukuda J, Tanaka T. Two instead of three embryo transfer in in-vitro fertilization. Hum Reprod 1995;10:2155–8.
5. Luke B, Keith L. The contribution of singletons, twins and triplets to low birth weight, infant mortality and handicap in the United States. J Reprod Med 1992;37:661–6.
6. Luke B, Bigger HR, Leurgans S, Sietsema D. The cost of prematurity: a case–control study of twins vs singletons. Am J Public Health 1996;86:809–14.
7. Spellacy WN, Handler H, Feare CD. A case–control study of 1253 twin pregnancies from a 1982–1987 perinatal data base. Obstet Gynecol 1990;75:168–71.
8. Martin JA, Hamilton BE, Sutton PD, Ventura SJ, Menacker F, Kirmeyer S. Births: final data for 2004. Natl Vital Stat Rep 2006;55(1):1–104.
9. The European IVF-monitoring programme (EIM) for the European Society of Human Reproduction and Embryology (ESHRE), Andersen AN, Gianaroli L, Felberbaum R, de Mouzon J, Nygren KG. Assisted reproductive technology in Europe, 2002. Results generated from European registers by ESHRE. Hum Reprod 2006;21:1680–97.
10. Centers for Disease Control and Prevention. 2002 Assisted reproductive technology success rates: national summary and fertility clinic reports. Atlanta, GA: Centers for Disease Control and Prevention, 2004:20–71.
11. Practice Committee of the Society for Assisted Reproductive Technology. Practice Committee of the American Society for Reproductive Medicine. Guidelines on number of embryos transferred. Fertil Steril 2006;86(Suppl 5):S51–2.
12. Gleicher N, Campbell DP, Chan CL, Karande V, Rao R, Balin M, et al. The desire for multiple births in couples with infertility problems contradicts present practice patterns. Hum Reprod 1995;10:1079–84.