Incomplete and inconsistent reporting of maternal and fetal outcomes in infertility treatment trials

Pregnancy outcomes and adverse outcomes in infertility trials are reported to varying extents; for example, 35% of clinical trials reported no information on pregnancy loss, only 43% reported adverse events during the preconception treatment period, and only 7% reported any serious adverse events. Incomplete reporting limits the value of these studies in counseling patients on the risk-benefit ratio of treatment to themselves and their infants. (Fertil Steril® 2011;95:2527–30. ©2011 by American Society for Reproductive Medicine.)

Key Words: Miscarriage, multiple pregnancy, morbidity, mortality, fetal anomaly, intrauterine fetal demise, adverse event

There is no standard for reporting outcomes of infertility trials. Although the ultimate cure for infertility is a healthy infant delivered by a healthy mother (1), it is often difficult to extract information about maternal and fetal risks from published trials. Infertility trials are unique in that they potentially involve three subjects: mother, father, and fetus or infant, often cared for by separate physicians. Two of these (mother and fetus) fall into vulnerable categories as defined by federal guidelines. Current standards for reporting randomized clinical trials, such as the Consolidated Standards of Reporting Trials (CONSORT) guidelines, have not been modified to address these unique issues of infertility trials (2). Many groups have advocated tracking the singleton live birth rate as the primary outcome of infertility trials (1, 3, 4). Most infertility trials are underpowered to detect both differences in the singleton live birth rate and rare adverse events to the mother or child. Small trials can be integrated into a systematic review or meta-analysis where a more accurate risk-benefit ratio can be determined, and patterns of rare adverse events to mother and fetus can be assessed.

Lisa Dapuzzo, M.D.[a]
Faith E. Seitz, B.S.[b]
William C. Dodson, M.D.[b]
Christina Stetter, B.S.[c]
Allen R. Kunselman, M.A.[c]
Richard S. Legro, M.D.[b]

[a] *Department of Obstetrics and Gynecology, Lehigh Valley Hospital, Allentown, Pennsylvania*
[b] *Department of Obstetrics and Gynecology, Penn State College of Medicine, Hershey, Pennsylvania*
[c] *Department of Public Health Sciences, Penn State College of Medicine, Hershey, Pennsylvania*

Received March 25, 2010; revised and accepted February 21, 2011; published online March 24, 2011.
L.D. has nothing to disclose. F.E.S. has nothing to disclose. W.C.D. has nothing to disclose. C.S. has nothing to disclose. A.R.K. has nothing to disclose. R.S.L. has nothing to disclose.
Supported in part by Public Health Service grants U54 HD34449, U10 HD 38992, and RO1HD433332.
Reprint requests: Richard S. Legro, M.D., Department of Obstetrics and Gynecology, Pennsylvania State University College of Medicine, 500 University Drive, Hershey, PA 17033 (E-mail: RSL1@psu.edu).

We sought to determine what outcomes were regularly reported in the recent past in the leading journals in obstetrics and gynecology, as they tend to set the standard and can serve as a "best practice" example. We performed a literature review of all clinical and randomized controlled trials in infertility from April 1, 2004, to March 31, 2010, using PubMed and online archives. Inclusion of articles in this review had to have all of the following components: [1] clinical or randomized controlled trial; [2] human subjects, English language, adults $\geq$19 years of age either male or female, with documented infertility; [3] intervention with pregnancy as one of the outcomes; and [4] published in a top 10 journal in obstetrics and gynecology from 2008 Institute for Scientific Information rankings.

In the absence of uniform definitions, conception, pregnancy, and live birth outcomes were analyzed according to criteria reported from the Pregnancy in Polycystic Ovary Syndrome (PPCOS) clinical trial as our standard of comparison (5). In the PPCOS trial, conception was defined as any positive serum hCG level, a pregnancy was defined as an intrauterine pregnancy with fetal heart motion as determined by transvaginal ultrasound examination, and live birth was defined as delivery of any viable infant. Further intrauterine pregnancies were stratified by singleton, twin, triplet, or other. Pregnancy loss was recorded as total loss among subjects who conceived but also reported separately as a first-trimester loss, or a second- or third-trimester loss. The mode of first-trimester loss was further characterized as either an ectopic pregnancy, biochemical pregnancy, or intrauterine pregnancy without fetal heart motion on ultrasound scan. Pregnancy loss after confirmed fetal heart motion was reported as a separate category.

Adverse events in the reviewed trials were reported by whether they occurred before conception, after conception, or during delivery, and serious adverse events were reviewed to see whether they resulted in maternal or neonatal morbidity and mortality. The U.S. Food and Drug Administration defines serious adverse events as any event in a clinical trial that [1] results in death, [2] is life threatening, [3] requires inpatient hospitalization or prolongation of

existing hospitalization, [4] results in persistent or significant disability or incapacity, [5] is a congenital anomaly or birth defect, or [6] requires intervention to prevent permanent impairment or damage. We did not obtain Institutional Review Board approval because this research did not involve any human subjects; it was a bibliometric review.

The literature review of all clinical trials in infertility meeting our prespecified criteria resulted in a total of 373 articles of which 294 met our inclusion criteria on closer review (Fig. 1). Reasons for exclusion of articles included that pregnancy was not an outcome, the study was a cohort study and not randomized, and there was no documented infertility before intervention. Most of the articles provided a power analysis for their study (213/294 or 73%). The mean number (±SD) of participants per study was 225 ± 311 (range 8–3,371) and of pregnancies was 76 ± 112 (range 2–1,229). The pregnancy rate could be identified in 96% (282/294) of studies (Fig. 2). Only 73% (216/294) of the studies had a specific objective measurement to determine pregnancy, such as hCG level, identifiable gestational sac on ultrasound scan, or positive fetal heart motion by ultrasound. Conception rate, defined as positive pregnancy test according to our standard, was reported less commonly than pregnancy rate. First-trimester twin rate was reported in 29% (84/294) of articles and triplets in 18% (53/294) of articles. The majority of studies did not report live birth rates or multiple live birth rates. Thirty-five percent of articles did not note whether there were any pregnancy losses. In some adverse events were reported in the text and in others simply in table format. Only 43% of articles reported adverse events during the preconception treatment period. Adverse events were reported at a similar rate during the postconception period; however, the rate of reporting dropped sharply for intrapartum and postpartum events. Neonatal morbidity and mortality are only rarely reported (8% or 23/294).

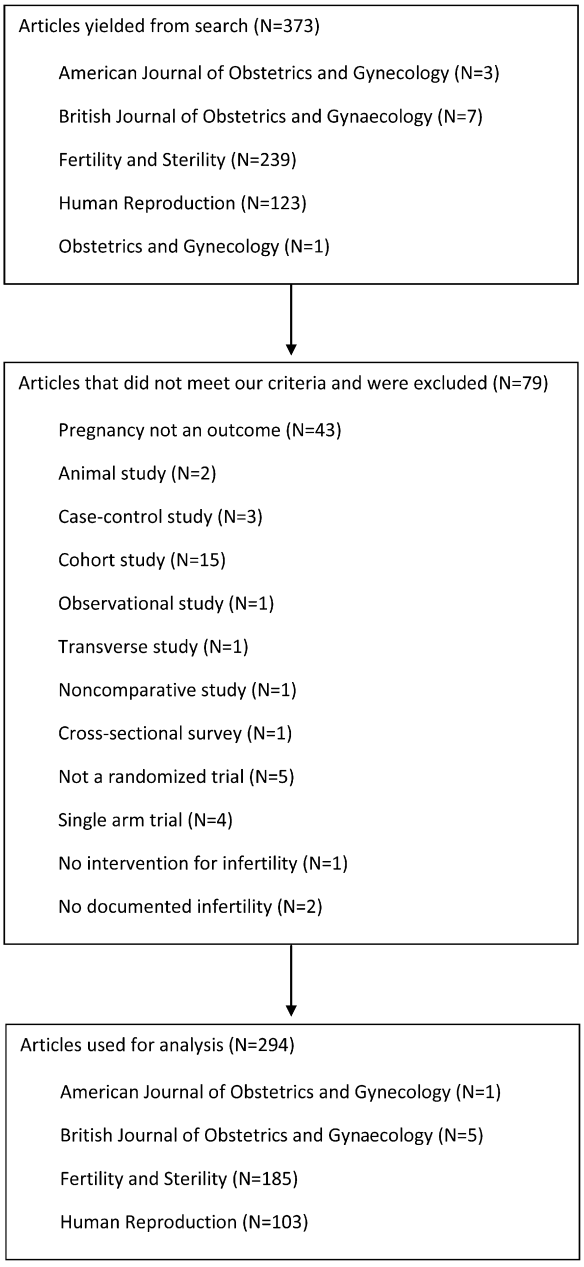
Of the 294 articles reviewed, 21 articles (7.1%) specifically used "serious adverse event" terminology to report whether or not they had such events. Thirteen of these 21 articles stated that there were no serious adverse events reported, whereas the other 8 articles reported on a variety of maternal and fetal complications (6–13) (Table 1). We found 30 other articles that, although they did not specifically use the words "serious adverse event," reported on death, hospitalization, congenital malformation, or in utero fetal demise (we searched specifically for these terms). All deaths were fetal; we found no cases of maternal mortality.

In addition to the fetal anomalies noted in the article by Hagemann et al. (6), our search of specific terms revealed only 10 other articles that reported on the presence or absence of congenital anomalies. Some only cited a rate or referred only to major malformations. Other specific anomalies reported included hypospadias, cleft palate, congenital heart defects, and limb-body complex. Additionally there were four elective terminations and one elective reduction, because of fetal abnormalities detected with amniocentesis or ultrasound screening. There was also one infant with Down syndrome.

Our review of clinical trials of infertility treatment over a 6-year period demonstrated large gaps and variability in the reporting of outcomes and events of interest to clinicians and patients. Although almost all articles reported on pregnancy rates, the definition of pregnancy varied from study to study. Few studies followed subjects to term to determine live birth rates, which ultimately is the primary goal of infertility therapy. Pregnancy loss was not mentioned in one third of studies, and, for those that did report it, it was difficult to determine the age of gestation at which the loss occurred and the type of loss. Multiple pregnancy, both as an outcome and as an adverse event, often is not identified clearly. Finally, adverse events during the preconception treatment period and after conception were reported in approximately 43% of articles, but fetal malformations and neonatal complications are reported infrequently. Most clinical trials did not report the presence or absence of adverse events. On a positive note, almost three quarters of the

FIGURE 1

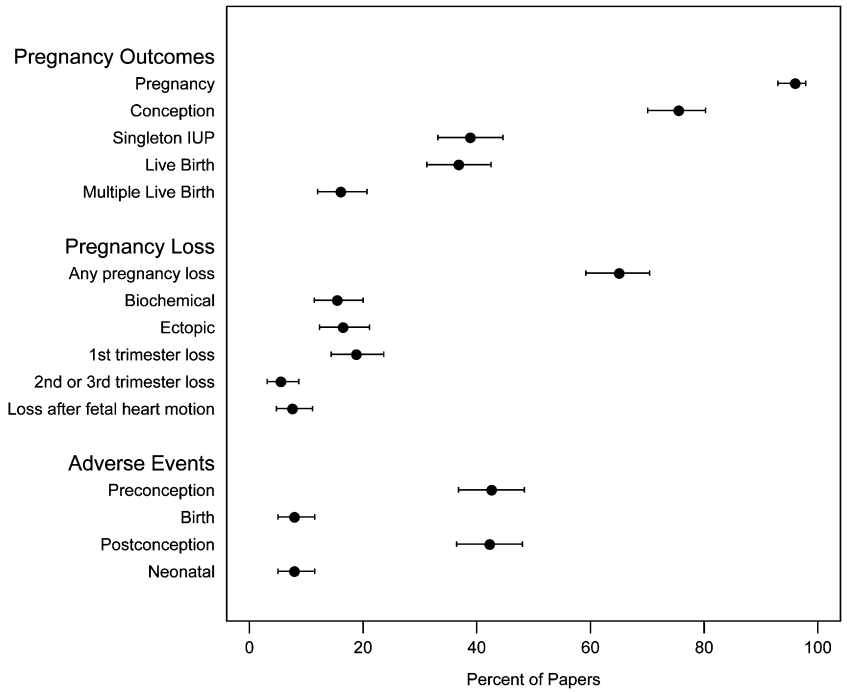
Flow chart of articles retrieved from the top impact journals in obstetrics and gynecology for this review.



Dapuzzo. Outcomes and infertility trials. Fertil Steril 2011.

FIGURE 2

The proportion of articles reporting the events of interest, for example, pregnancy and live birth, multiple pregnancy, adverse events, was calculated with exact 95% confidence limits. Conception was defined as any positive serum hCG level, a pregnancy was defined as an intrauterine pregnancy with fetal heart motion as determined by transvaginal ultrasound scan, and live birth was defined as delivery of any viable infant. IUP = intrauterine pregnancy.



Dapuzzo. Outcomes and infertility trials. Fertil Steril 2011.

trials performed some sort of power analysis to justify their sample size, although the chosen outcome was rarely live birth.

These data highlight the variable amount of relevant information that can be extracted from infertility trials to provide patients with an honest assessment of the risk-benefit ratio for infertility treatments. Success is defined variably, and pregnancy is without a standard definition. Our results are supported by a recent meta-analysis of randomized clinical trials in reproductive medicine abstracted from the Cochrane Database, which similarly noted that only few studies report live birth (only 22%) (14). The study using the Cochrane Database also noted that pregnancy rates (defined in the study as an intrauterine pregnancy at 6 weeks) were comparable with live birth rates in determining the effectiveness of a treatment (kappa value of 0.81; 95% confidence interval, 0.68–0.94) (14). Alternatively, perhaps 80% of trials did not report live birth rates because there were significant discrepancies between pregnancy and live birth rates, and publication bias favored pregnancy rates. Our data show that pregnancy loss, and importantly when in pregnancy it occurs, is rarely reported. A review of the quality of randomized controlled trials in infertility found that most failed to meet CONSORT guidelines and of those that did only 30% provided sufficient details to allow for an intention-to-treat analysis of the outcome "live birth" (15).

The effectiveness of treatment is only one aspect to consider when choosing a medical intervention. Patients also need to know the risks of infertility treatments. Multiple pregnancy remains perhaps the greatest risk of infertility treatment, and, as shown in our analysis, this is usually not reported. Serious adverse events are either rare or, on the basis of our data, underreported. Most articles that specifically used the term "serious adverse events" tended to be industry-sponsored trials, where events are collected prospectively and strictly categorized. Our data suggest that serious adverse events, when so defined and collected, are actually quite common in infertility trials. Birth defects remain rare and are rarely reported but are major concerns especially with newer fertility agents, such as letrozole (16). Without systematic collection and obligatory reporting of them in infertility trials, any adverse association is likely to be missed, as the story of thalidomide well illustrates.

Our study has weaknesses. Our standard for reporting as noted above is arbitrary. Another potential weakness is our failure to perform a systematic review. This might have yielded alternate results and conclusions. We chose journals that were highly reputable on the basis of impact factor and reasoned that the inclusion of lesser journals would only further drag down the results. Infrequent reporting of such events as multiple pregnancies, fetal malformations, or neonatal problems may occur because these actually are rare events. This remains an assumption unless such outcomes are routinely tracked and reported (including the acknowledgment of no adverse outcomes).

Our data indicate that pregnancy outcomes, adverse events, and fetal and neonatal outcomes are reported incompletely. This limits

TABLE 1

List of articles reporting "serious adverse events" including number and type of event.

Reference	No. of subjects in study	Total no. SAE	Types of SAE
Keye et al., 2004 (13)	110	8	n = 4, OHSS n = 2, ectopic pregnancy n = 1, subclavian venous thrombosis n = 1, pelvic pain
Kleinstein et al., 2005 (11)	430	1	n = 1, jugular vein thrombosis
Simons et al., 2005 (12)	178	2	n = 1, pelvic inflammatory disease n = 1, pelvic pain
Wilcox et al., 2005 (9)	185	7	n = 3, ectopic pregnancies n = 1, postprocedural hemorrhage n = 1, myasthenia gravis n = 1, OHSS Other not specified
Amer et al., 2009 (10)	72	1	n = 1, OHSS
Brinsden et al., 2009 (8)	150	10	Not specified
Devroey et al., 2009 (7)	1506	74	n = 23, OHSS n = 16, ruptured ectopic pregnancies Others not specified
Hagemann et al., 2010 (6)	121	6	n = 4, fetal anomalies n = 2, maternal pregnancy complications

Note: OHSS = ovarian hyperstimulation syndrome; SAE = serious adverse event.

Dapuzzo. Outcomes and infertility trials. Fertil Steril 2011.

our ability to design larger trials focused on live birth or avoidance of rare events. Perhaps more important, it deprives us of important information to inform patients of the risks and benefits of treatment, which could be assessed through periodic review and meta-analysis of the published trials. The development of standardized guidelines for reporting outcomes of infertility trials will improve the quality of published trials and ultimately clinical care.

REFERENCES

1. Legro RS, Myers E. Surrogate end-points or primary outcomes in clinical trials in women with polycystic ovary syndrome? Hum Reprod 2004;19:1697–704.
2. Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. Lancet 2001;357:1191–4.
3. Johnson NP. No more surrogate end-points in randomised trials: The PCOSMIC trial protocol for women with polycystic ovary syndrome using metformin for infertility with clomiphene. Aust N Z J Obstet Gynaecol 2006;46:141–5.
4. Min JK, Breheny SA, MacLachlan V, Healy DL. What is the most relevant standard of success in assisted reproduction? The singleton, term gestation, live birth rate per cycle initiated: the BESST endpoint for assisted reproduction. Hum Reprod 2004;19:3–7.
5. Legro RS, Barnhart HX, Schlaff WD, Carr BR, Diamond MP, Carson SA, et al. Clomiphene, metformin, or both for infertility in the polycystic ovary syndrome. N Engl J Med 2007;356:551–66.
6. Hagemann AR, Lanzendorf SE, Jungheim ES, Chang AS, Ratts VS, Odem RR. A prospective, randomized, double-blinded study of assisted hatching in women younger than 38 years undergoing in vitro fertilization. Fertil Steril 2010;93:586–91.
7. Devroey P, Boostanfar R, Koper NP, Mannaerts BM, Ijzerman-Boon PC, Fauser BC. A double-blind, non-inferiority RCT comparing corifollitropin alfa and recombinant FSH during the first seven days of ovarian stimulation using a GnRH antagonist protocol. Hum Reprod 2009;24:3063–72.
8. Brinsden PR, Alam V, de Moustier B, Engrand P. Recombinant human leukemia inhibitory factor does not improve implantation and pregnancy outcomes after assisted reproductive techniques in women with recurrent unexplained implantation failure. Fertil Steril 2009;91:1445–7.
9. Wilcox J, Potter D, Moore M, Ferrande L, Kelly E. Prospective, randomized trial comparing cetrorelix acetate and ganirelix acetate in a programmed, flexible protocol for premature luteinizing hormone surge prevention in assisted reproductive technologies. Fertil Steril 2005;84:108–17.
10. Amer SA, Li TC, Metwally M, Emarh M, Ledger WL. Randomized controlled trial comparing laparoscopic ovarian diathermy with clomiphene citrate as a first-line method of ovulation induction in women with polycystic ovary syndrome. Hum Reprod 2009;24:219–25.
11. Kleinstein J. Efficacy and tolerability of vaginal progesterone capsules (Utrogest 200) compared with progesterone gel (Crinone 8%) for luteal phase support during assisted reproduction. Fertil Steril 2005;83:1641–9.
12. Simons AH, Roelofs HJ, Schmoutziguer AP, Roozenburg BJ, van't Hof-van den Brink EP, Schoonderwoerd SA. Early cessation of triptorelin in in vitro fertilization: a double-blind, randomized study. Fertil Steril 2005;83:889–96.
13. Keye WR Jr, Marrs RP, Check JH, Schnell V, Surrey M, Marshall DC. Evaluation of mixed protocols with Bravelle (human-derived FSH) and Repronex (hMG) to assess clinical efficacy (EMBRACE) in women undergoing in vitro fertilization. Fertil Steril 2004;82:348–57.
14. Clarke JF, van Rumste MM, Farquhar CM, Johnson NP, Mol BW, Herbison P. Measuring outcomes in fertility trials: can we rely on clinical pregnancy rates? Fertil Steril 2010;94:1647–51.
15. Dias S, McNamee R, Vail A. Evidence of improving quality of reporting of randomized controlled trials in subfertility. Hum Reprod 2006;21:2617–27.
16. Tulandi T, Martin J, Al-Fadhli R, Kabli N, Forman R, Hitkari J, et al. Congenital malformations among 911 newborns conceived after infertility treatment with letrozole or clomiphene citrate. Fertil Steril 2006;85:1761–5.