# Live birth is the correct outcome for clinical trials evaluating therapy for the infertile couple

Kurt T. Barnhart, M.D., M.S.C.E.

Division of Reproductive Endocrinology and Infertility, Department of Obstetrics and Gynecology, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania

Well-designed and -conducted clinical trials are needed to further advance the field for reproductive medicine. However, current reporting of outcomes of trials is ambiguous and disparate. In this review it is offered that the preferred outcome for clinical trials in reproductive medicine should be live birth. Multiple births should be listed, and it should be specified whether this is multiple births per couple or multiple births per conception. The unit of measure should be women (or couples) and not cycles. The duration of exposure should also be clearly identified (i.e., treatment was one cycle, a prespecified number of cycles, or a period of time). Pregnancy loss should be specified, and the denominator should be those who conceived. Although live birth is the primary outcome, complications should be defined and reported, including multiple births and other objective markers, such as preterm delivery, small-for-gestational age, or stillbirth. (Fertil Steril® 2014;101:1205–8. ©2014 by American Society for Reproductive Medicine.)

**Key Words:** Infertility, clinical trial, outcome, live birth

**Discuss:** You can discuss this article with its authors and with other ASRM members at http://fertstertforum.com/barnhartk-live-birth-infertility-therapy/

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.

The goal of clinical research is to inform the science and to aid clinical medicine. Reproductive medicine has enjoyed tremendous growth both scientifically and clinically in the past decades. To further inform clinical care and to practice evidence-based medicine, there have been appropriate and widespread calls for well-conducted clinical trials in infertility (1, 2). However, clinical trials in infertility are challenging to conduct, and reporting has been incomplete and inconsistent (1–3). In the attempt to improve the transparency and impact of clinical trials, the Consolidated Standards of Reporting Trials (4) statement has been reviewed by experts in reproductive medicine. Consensus was sought and obtained to provide guidance for the specificity needed for reporting outcomes of clinical trials. One of the main deficiencies noted by this group was the lack of consistency in the reporting of the primary outcome in trials designed to improve fertility.

The goal of therapy in clinical reproductive medicine is to assist couples in starting or extending their families. How could this outcome be ambiguous? It should be relatively easy to just count our successes and report them. Should not this problem be as simple as counting the number of children in a family? It is obvious if a childless family now has a child. It should also be obvious if a family has grown from one child to two or more. The first stumbling block in this simple strategy is that not every pregnancy results in a single birth. Of course infertility trials must account for multiple births. However, the problem goes beyond multiple births: our literature is replete with disparate and confusing outcomes. Reported outcomes include stimulation parameters, number and quality of gametes (eggs and sperm), embryo survival rate, implantation rate, chemical pregnancy, clinical pregnancy, ongoing pregnancy, and live birth. At times pregnancy is further subdivided into twins, triplets, higher-order multiple gestations, vanishing twins, and vanishing triplets.

Confounding this difficulty is that some of these commonly reported terms are not uniformly defined. For example, is a chemical pregnancy the earliest form of pregnancy (and thus a positive outcome), or is it the earliest form of a miscarriage (and thus a poor outcome)? Success with IVF is often reported as the number of women with positive pregnancy test results. However, no couple is happy with the outcome of a chemical pregnancy (a pregnancy

Received January 22, 2014; revised February 24, 2014; accepted March 13, 2014.
K.T.B. has nothing to disclose.
Reprint requests: Kurt T. Barnhart, M.D., M.S.C.E., Division of Reproductive Endocrinology and Infertility, Department of Obstetrics and Gynecology, Perelman School of Medicine at the University of Pennsylvania, 3701 Market Street, Suite 800, Philadelphia, Pennsylvania 19104 (E-mail: kbarnhart@obgyn.upenn.edu).

Fertility and Sterility® Vol. 101, No. 5, May 2014 0015-0282/$36.00

http://dx.doi.org/10.1016/j.fertnstert.2014.03.026

loss). Clinical pregnancy is often defined as a gestational sac identified with ultrasound, but when that sac is identified is not uniform. It can be as early as 6 weeks or perhaps as late as 12 weeks. Thus, a clinical pregnancy in one trial may be a chemical pregnancy in another. The term "ongoing pregnancy," which is meant to suggest that there is a very high likelihood of a pregnancy continuing to term, is equally ill defined. At times ongoing pregnancy is classified as a pregnancy that has fetal cardiac activity at 8 weeks, 10 weeks, or 12 weeks, or often it is not specified.

Live birth, the preferred primary outcome, is used in some fertility trials. However, even then the definition of a live birth is not without controversy. The Society for Assisted Reproductive Technology database defines live-birth delivery as birth of one or more live-born infants (with no specificity of gestational age), with delivery of multiple infants counted as one live-birth delivery. A multiple birth is defined as a birth of two or more infants, at least one of whom was live-born. By contrast, the Centers for Disease Control and Prevention's National Center for Health Statistics, which bases its statistics on live birth records rather than delivery records, classifies the delivery of a single live-born infant with one or more stillbirths as a singleton birth (5). How far does a gestation need to progress to be considered a live birth? For example, if a fetus is born at 19 weeks, with cardiac activity and respiratory effort, is it a live birth? Some have suggested that the definition of a live birth should only be in what is considered a viable gestational age, such as 23 or 24 weeks. We propose a definition of a live birth to include a gestational age of $\geq$20 weeks, reflecting the World Health Organization definition (6).

Why is live birth not universally accepted as the primary outcome in infertility trials? Many excuses have been proffered, including that data are not easy to get because of the fragmentation of reproductive and obstetric care. Once a clinical or ongoing pregnancy is identified, the mother is referred to another practice—an obstetrician, a midwife, or at times even a completely different institution, possibly in a different state or region. Therefore, collection of these data to a clinical investigator in reproductive medicine is a "burden." The excuse should be dismissed simply as lazy. It is understood that the conduct of a clinical trial is expensive and burdensome. However, the small incremental cost is necessary to obtain the appropriate outcome. If one is to conduct appropriate high-quality clinical research, then the cost to obtain information on the circumstance of birth after intervention is a necessity and not a luxury. Randomized clinical research should not be performed on the cheap.

Some have suggested that information gleaned from the clinical trial is so time-sensitive that one cannot wait the additional 7 months necessary to find out whether a pregnancy conceived results in a live birth. Clearly, if the results are meaningful they are worth the wait. It is also important to understand and report the perinatal outcomes experienced by mother and child (7).

It is possible that pregnancy and live birth as endpoints in a clinical trial are "comparable." This was objectively assessed by Clarke et al. (8), who noted that only 20% of 654 randomized clinical trials reported both live birth and clinical pregnancy as an outcome. The loss from clinical pregnancy to live birth was approximately 19%. The differential absolute loss in those with an active therapy compared with those who conceived without medical intervention (controls) was similar at approximately a 5.4% and 5.5% loss rate (30.3% vs. 24.9% in those with treatment, 27.9% vs. 22.4% in controls). Therefore, one possible conclusion is that the 19% pregnancy loss can simply be extrapolated from the clinical pregnancy to achieve a reasonable proximity of the live birth (8). Thus, clinical pregnancy can be a surrogate marker for live birth. However, it is possible that an intervention may have a differential effect on miscarriage and/or survival of a pregnancy compared with an unassisted pregnancy. This can only be noted if clinical pregnancy and live births are reported. One needs a very strong rationale to accept a surrogate endpoint when the true clinical endpoint can easily be obtained (9, 10).

There are many examples in medicine in which surrogate endpoints have misled. For example, some medications decreased arrhythmias but paradoxically increased the risk of death from other causes (11). Fluorides increase bone mineral density in women with osteoporosis but lead to more fractures (12). This important phenomenon is pertinent in reproductive medicine as well. For example, live birth is higher with fresh compared with frozen transfer (13), but perinatal outcomes for children seem to be worse when a child is conceived with fresh transfer (compared with a frozen–thawed transfer) (14–16).

The reporting of standardized secondary outcomes is also important in infertility trials. If a conception does not result in a live birth, the outcome of that pregnancy should be reported. Timing of the loss is important and should be reported. Is the pregnancy loss in the first trimester, the second trimester, or did it result in a stillbirth? If pregnancy is never visualized, there are consensus documents that can be used to classify the ultimate clinical outcome, which may include a visualized or nonvisualized ectopic pregnancy, a histologic intrauterine pregnancy, or a resolved or treated pregnancy of unknown location (17). The outcome of all gestations should be reported, including spontaneous and active fetal reduction.

Another potential difficulty in reporting outcomes in fertility trials is the unit of measurement. If the unit randomized is something other than the woman (or the couple), results can be misleading. Randomization of eggs, embryos, or cycles can result in a unit of analysis error (18). Eggs or embryos from the same women are interrelated, and when combined with those randomized from other women, the premise of independence necessary for statistical analysis can be violated. Second, multiple observations per woman can lead to an unpredictable bias in the estimate of a treatment difference. It will also exaggerate the apparent sample size, giving more precision to an outcome that may be inappropriate. Therefore, articles will have a spurious narrowing of confidence intervals and lower *P* values, potentially resulting in a type I statistical error (whereby an apparent statistical association is noted when in truth none occurs). Many reported trials had a "unit of analysis" error (18). Sometimes the actual endpoint is unclear, and at times pregnancy is not even reported (1). This can result in misinformation, and worse,