

REPRODUCTIVE HEALTH CARE POLICIES AROUND THE WORLD

Toward a More Meaningful In Vitro Fertilization Success Rate

RAYWAT DEONANDAN,^{1,3} M. KAREN CAMPBELL,¹ TRULS ØSTBYE,¹ and IAN TUMMON²

Submitted: January 29, 2000

Accepted: January 29, 2000

Purpose: The objective was to explore the variability in in vitro fertilization (IVF) success rates.

Methods: Published success rates from IVF clinics in North America were investigated to establish types of biases and potential inaccuracies.

Results: Success rates reported by IVF clinics vary with regard to the indices and patient populations used to compute them. Selection bias and misunderstood statistics are major factors contributing to the inappropriateness of certain rates.

Conclusions: The influence of privatization and market forces also may contribute to the need to oversimplify IVF statistics.

KEY WORDS: Bias; intracytoplasmic sperm injection; in vitro fertilization; statistics; success rate.

INTRODUCTION

The continuing popularity of in vitro fertilization (IVF) among common infertility therapies is complicated by the variable methods of calculating and re-

porting the procedure's success statistics. Table I lists some of the elements used in defining various success rates. Clinics have been accused of distorting their outcome rates (1, 2) in an attempt, unconscious or otherwise, to artificially inflate estimates of their programs' success. Since the success rate continues to be regarded as the primary criterion of a clinic's proficiency (2), it can be expected that the need to compete for patients in an open-market system may lead to a certain bias, subtle and perhaps unconscious, in the publishing of IVF outcomes.

A review of the nature of IVF success reporting is therefore indicated, building on the work of page *et al.* (3) and Wilcox *et al.* (4), so that a greater understanding of its complicated nature may be rendered, and that biases that may mislead a potential patient, clinician, administrator, or policymaker can be minimized.

FACTORS INFLUENCING SUCCESS RATES

Patient Population

There are inadvertent selection biases and other patient population characteristics that may lead to misestimation of success rates. The most likely selection bias occurring in IVF involves the rejection of patients with bad prognoses and the encouragement of those with good prognoses. A common error affecting the construction of an IVF success

¹ Department of Epidemiology and Biostatistics, University of Western Ontario, Ontario, Canada.

² Department of Gynaecology and Reproductive Medicine, London Health Sciences Center, London, England.

³ To whom correspondence should be addressed at Department of Epidemiology and Biostatistics, University of Western Ontario, London, Ontario N6A 5C1, Canada.

rate is the “selecting out” of patients who would otherwise reduce the success rate. Selection bias may overestimate success rates for an IVF program and affect its subsequent attraction of patients and policymakers.

Selection of Patients. One obvious form of selection bias arises from age cutoffs. If women above a certain age are denied treatment, then those clinics have ostensibly eliminated those potential patients with the most decreased fecundity and therefore the least chance of success. Waiting lists are another potential source of selection bias if patients with good prognoses are pushed to the front of the line, while those with poorer clinical expectations are encouraged to wait.

A more subtle selection bias arises from the change in patient population between cycles (2), where clinicians shape patients’ decisions via counseling. Although this is very appropriate clinical behavior, it also biases the success rate upward for those in later cycles. As well, this practice reduces the internal consistency of any success index that uses cycle number as its denominator. Here, the responsible application of policies intended to better serve the individual patient has generated both a selection bias and a statistical difficulty.

Selection bias is unabashedly manifested in the establishment of specialty IVF clinics: specialization of clinics leads to highly selected populations. Specialty clinics are for specific underlying infertility diagnoses and demographic types. Clinics that specialize in offering healthy donor gametes, for example, may benefit from higher success rates than those that attempt to work with patients’ own gametes. The system of cross-referrals allows for a higher concentration of patients with poorer prognoses in centers with less stringent admission criteria.

Conversion of IUI Cycles. The practice of converting hyperstimulated intrauterine insemination (IUI) patients to IVF treatment enriches a patient pool with individuals experiencing less serious infertility issues. In vitro fertilization may be offered to IUI patients midtreatment to control situations in which fertility drugs are putting these women at risk for multiple pregnancies or other complications; the IVF patients pool therefore is benefiting from women who at that moment are “hyperfertile.” Published success rates tend not to mention whether IUI services are offered at that clinic, or if IUI patients are converted to IVF cycles. For that matter, inclusion/exclusion criteria are rare in published IVF success reports.

Table I. Some Elements Commonly Found in IVF Success Rates

Numerators
Biochemical pregnancies
Clinical pregnancies
Number of gestational sacs
Live deliveries
Denominators
Ovarian stimulation procedures
Natural cycles of oocyte maturation
Oocyte retrieval procedures
Embryos transferred
Procedures per patient
Patients

Insurance Coverage. For an explicit comparison of public versus private clinics, the evaluation of selection bias is not straightforward. Though medical insurance guidelines cause public clinics to attract women with specific underlying diagnoses, it is unclear how these diagnoses affect the outcome of IVF therapy. Couples undergoing the private therapy of intracytoplasmic sperm injection (ICSI) for reasons of male infertility, for example, might be expected to have a higher chance of success by virtue of the woman’s better reproductive health relative to that of her partner.

This is relevant, for example, for clinics in the province of Ontario. Because ICSI is not funded by the Ontario Health Insurance Plan (OHIP), the subsequent enhancement of private clinics’ pool of ICSI patients may improve the rates of such clinics relative to their public counterparts. This argument may be simplistic, however, given the heterogeneous nature of clinics’ patients.

Garcia (5) points out that only a few American states (Arkansas, Connecticut, Hawaii, Illinois, Maryland, Massachusetts, Rhode Island, and Texas) have introduced bills requiring insurance companies to provide coverage for assisted reproduction procedures. With such limited and sporadic coverage, there exists a bias within IVF patient populations with respect to patient selection, pathology, and treatment. This bias is effected in a demographic gap between individual states that offer insurance, since the criteria for coverage differs, and also between insured states and uninsured states where the difference can be expected to be largely socioeconomic.

Statistical Reporting of Rates

Table I lists the elements that typically appear in the numerators and denominators of IVF success in-

dices. The most common and clinically convenient combination is the number of pregnancies per embryo transfer or per cycle. This reflects a concern for the cost of independent procedures since each cycle, reflecting both an oocyte retrieval procedure and an embryo transfer procedure, can cost a patient \$8,000 US (6).

Wilcox *et al.* (4) conducted a systematic review of chosen indices and were of the opinion that since the ultimate goal of IVF is the birth of a healthy living infant, the best numerator is probably the number of live deliveries, together with the frequency of multiple live infants per delivery. The number of live deliveries, however, is substantially less than the number of pregnancies in the same population, and therefore less impressive.

Indexing by Cycle Instead of by Patient. Using the treatment cycle as the denominator of the success index makes several inappropriate assumptions. Because in many ways IVF is a subjective therapy based in large part on a physician's personal appraisal of his patient's responses, an individual's treatment can be bettered in subsequent cycles as more is learned about her particular physiology. Cycles, then, are not statistically independent events and should not appear alone in an index denominator.

Furthermore, those who are going to become pregnant will likely be successful in their first cycle (7). Patients who need to undergo more procedures are those who are less likely to be successful in any given cycle. Measurement based on cycles therefore is biased toward those who are less successful at becoming pregnant. In that case, success rates should be stratified by cycle number.

The error of considering treatment attempts and not individual patients as the units of analysis affects more complex statistical inquiries such as risk factor modeling. This type of analysis typically involves the association of a variety of clinical or demographic risk factors with potential outcomes, such as whether or not the patient becomes pregnant or gives birth. In the IVF literature, the application of logistic or multiple regression analyses has been performed using cycle number and not the individual patient as the unit of analysis.

The results of these analyses are typically odds ratios or relative risks for a list of factors that potentially affect the outcome of the IVF therapy. Given the inappropriateness of using cycle number, such results and their therapeutic impacts are in question. The use of cycles in this respect constitutes a statistical error in independence. Since the aforementioned

analytical techniques incorrectly assume that knowledge of the events of one cycle do not confer knowledge of the events in subsequent cycles, the conclusions of those techniques are compromised. The direction and magnitude of the error in computing odds ratios and relative risks are dependent on the proportion of patients who are successful in earlier versus later cycles.

From the perspective of the patient, it is most desirable to know the chances of success based on the number of patients who have gone through the program and not on the total number of attempts made by all patients. An administrator may wish an index to be dependent on the total number of cycles for accounting and administrative reasons, and a clinician needs to know the projected success of an individual undergoing a varying number of cycles. A compromise worthy of further investigation may be the computation of the mathematical product "patient-cycles" as a denominator reflective of both the extent of treatment and the number of individuals having undergone it, thus allowing cycle number to function as more of a weighting factor than as a unit of analysis per se.

Heterogeneity of Population and Clinical Guidelines. An obvious issue that arises from a dependence on simple indices is that they all imply a homogeneous patient population. In truth, a consumer wishing to evaluate a clinic for her potential treatment needs to consider that clinic's history of treating patients similar to herself in terms of underlying etiology of infertility, age, parity, and type of intervention. The simple indices are easily cited and understood, and thus remain the reporting styles of choice.

Inappropriate Use of Life Table Analysis. Life table analysis has been used to estimate the chances of "succeeding" at IVF. The underlying rationale is the modeling of the IVF process in such a way that pregnancy or live birth are the desired outcomes that represent end points of the life table analysis; failure to reach the end point within the expected time frame is considered a "loss to follow-up." This is a statistically problematic approach for a number of reasons, not the least of which is that life table analysis is meant to follow the survival of a homogeneous population after a single event, while IVF constitutes a series of interventions that may vary in timing, extent, and detail from patient to patient.

Furthermore, the IVF situation violates an important assumption of life table analysis, namely that any group of patients who are lost to follow-up experience the same chances of "success" as those who

remain in the program. As has been mentioned, chances of “success” are greater in earlier cycles, so the success chances of individuals lost to follow-up depends at least in part on when they left the program.

A further statistical assumption of this analysis is that individuals who are indeed lost to follow-up are lost or drop out of the program for reasons independent of the treatment regimen. In IVF, however, patients are typically lost to follow-up specifically because the treatment has failed them and clinical judgment has predicted that their chance of success is low.

Wilcox *et al.* (4) suggest an improvement of this analysis using a two-parameter approach. By obtaining estimates of the chances of an IVF patient becoming pregnant each month and an estimate of the proportion of patients who will never become pregnant, the authors imply that a suitable modification may be made to the loss-to-follow-up assumptions.

Other Outcomes

The simple indices used in Table II fail to reflect

any information about risks to the patient’s or offspring’s short-term or long-term health. If a given therapy reports a 34% chance of “success,” for instance, does this imply a 66% chance of “failure?” The answer would depend, of course, on one’s definition of failure: the inability to obtain a child or a more complex combination of various unwanted outcomes.

The use of fertility-enhancing drugs commonly used in IVF sometimes carries with it unpleasant side effects including nausea, hypertension, and rarely allergic reactions (e.g., see ref. 8). Furthermore, patients undergoing IVF may suffer from a high number of multiple and ectopic pregnancies, miscarriages, and, as would be expected from this specific patient group, a comparatively higher number of anomalous birth events relative to the naturally conceiving population (9–11). If an IVF pregnancy is achieved at all, there is evidence to suggest that that pregnancy has less of a chance of being carried to term than would a natural pregnancy (12). The negative psychosocial, medical, and financial consequences of such events are rarely explored.

Hallam Medical Centre’s use of a cumulative conception rate (see Table II) results in a high success rate but reflects the experiences of five cycles of pa-

Table II. Methods of Success Reporting for Selected Clinics

Clinic	Period	Success rate	Nature of measure	Source
Ottawa Civic Hospital	1994	0.25	Live child/embryo transfer	Civic Hospital (17)
Ottawa Civic Hospital	1994	0.35	Pregnancy/embryo transfer	Civic Hospital (17)
London Health Sciences Centre	1995	0.20	Live birth/fresh transfer cycle	London Health Sciences Centre (18)
London Health Science Centre	1995	0.26	Pregnancy/fresh embryo transfer	London Health Sciences Centre (18)
University Hospital (now London Health Sciences Centre)	1984–1987	0.261	Pregnancy/couple treated	Yuzpe <i>et al.</i> (19)
Sahlgrenska Hospital, Sweden	1990–1992	0.28	Number of women with deliveries/cycle (only for women having completed 4th cycle)	Bergh <i>et al.</i> (20)
Sahlgrenska Hospital, Sweden	1990–1992	0.276	Number of women with deliveries/cycle (only for women who have completed 1st cycle)	Bergh <i>et al.</i> (20)
Sahlgrenska Hospital, Sweden	1990–1992	0.513	Cumulative number of women with deliveries/couples commencing treatment	Bergh <i>et al.</i> (20)
Hallam Medical Centre, UK	1984–1991	0.498	Cumulative conception rate (after 5 cycles)	Tan <i>et al.</i> (21)
Clinics across US	1993	0.183	Deliveries/cycle	ASRM (22)
Foothills Hospital, Calgary	1985	0.086	Clinical pregnancies/100 couples accepted	Taylor <i>et al.</i> (23)
Hallam Medical Centre, UK	1984–1991	0.39	Cumulative live birth rate (after 5 cycles)	Tan <i>et al.</i> (21)

tients, implying that the patients have undergone both the positive and the negative aspects of IVF therapy up to five times. There is no way to consider separate aspects, either qualitatively or quantitatively, from the 50% rate reported by Hallam Medical Centre.

The numbers listed for Ottawa Civic Hospital are closer to the rates researchers have come to expect in IVF, but the difference between 25% and 35% is quite large, exemplifying the differing importance of the two chosen end points, live birth or pregnancy. It could be argued that knowledge of both such indices is important to a patient/consumer, since together they reflect the program's ability to bring a conception to a desirable conclusion.

DISCUSSION

At present, IVF is one of the few therapies in Canada that exists both as privately and publicly funded enterprises. In Ontario, the growth of private IVF clinics has become most noticeable after 1994, when government policy changed with regard to which conditions warranted full OHIP funding (13). Among the underlying diagnoses of infertility that qualify for OHIP funding are bilateral tubal occlusion and infertility caused by idiopathic disease, while more common problems such as male-factor infertility and age-related indications for IVF go unfunded. Patients not fulfilling the "fundable" criteria may be referred to a growing number of private IVF clinics.

Clinics in the United States and to a lesser extent in Europe have always been operated privately and their experiences may prove appropriate for predicting similar trends in Canada as IVF privatization increases. Most American and some Canadian assisted reproduction clinics report their results to a standardizing body—the American Society for Reproductive Medicine (ASRM)—which collates and publishes the results yearly en masse using objective standards.

The ASRM reports suffer, however, from the same limitations listed above and so offer little in the way of decisive data for an administrator or potential consumer. In short, the published ASRM data does not permit valid clinic-to-clinic comparisons. To its credit, the 1995 ASRM report (14) represents the first collation of such data to address the long-standing criticism of indexing by cycle rather than by patient. An important first step toward reducing selection biases

would be the compilation of such data categorically by inclusion/exclusion criteria. Knowledge of such criteria would help to identify those patients treated by specialty clinics (those geared to certain etiologies of infertility, for example) and those who perhaps have been referred to less restrictive clinics from clinics displaying exceptionally good success statistics.

It should be pointed out that few computations can reflect the degree of emotional well-being that an assisted reproduction intervention may confer, whether or not that intervention is successful. Simply having sought treatment for one's infertility may offer its own psychological reward (15,16). Thus, any purely mathematical success rate would be an underestimation in that respect. Nevertheless, a greater need for detail and clarity is unavoidably indicated; no single index is sufficient for representing a program's achievements or philosophies.

Discrepancies in IVF success reporting is an acknowledged phenomenon, though solutions are rarely proposed. Speculations of its underlying motivation are even rarer, possibly because such conjectures venture into the realm of market forces, an uncomfortable area for many medical researchers. It is nonetheless conceivable that the need to advertise for patients in a privatized IVF setting contributes in part to the phenomenon of vague, ambiguous, or minimalist success reporting.

REFERENCES

1. Soules MR: The in vitro fertilization pregnancy rate: Let's be honest with one another. *Fertil Steril* 1985;43:511–513
2. Hershlag A, Kaplan EH, Loy RA, DeCherney AH, Lavy G: Selection bias in in vitro fertilization programs. *Am J Obstet Gynecol* 1992;166:1–3
3. Page EW, Villee CA, Villee DB: *Human Reproduction—Essentials of Reproductive and Perinatal Medicine*. Toronto, WB Saunders Co., 1981
4. Wilcox LS, Peterson HB, Haseltine FP, Martin MC: Defining and interpreting pregnancy success rates for in vitro fertilization. *Fertil Steril* 1993;60:18–25
5. Garcia JE: Profiling assisted reproductive technology: The Society for Assisted Reproductive Technology registry and the rising costs of assisted reproductive technology. *Fertil Steril* 1998;69:624–626
6. Seibel MM, Bernstein J: Cost-effective trial balloons and the potential for bias: The case against in vitro fertilization. *J Women's Health* 1994;3:411–414
7. Page H: Calculating the effectiveness of in vitro fertilization. A review. *Br J Obstet Gynecol* 1989;96:334–339
8. Scoccia B, Blumenthal P, Wagner C, Prins G, Scommegna A, Marut E: Comparison of urinary human follicle-stimulating hormone and human menopausal gonadotropins for ovarian

- stimulation in an in vitro fertilization program. *Fertil Steril* 1987;48:446–449
9. Corabian P, Hailey D: The efficacy and adverse effects of in vitro fertilization and embryo transfer. *Int J Technol Assess Health Care* 1999;15:66–85
 10. Yu V: Assisted reproduction technology, multiple births, and adverse perinatal outcome. *Croatian Med J* 1998;39:208–211
 11. Porreco R, Schoolcraft C, Schoolcraft W: Pregnancy outcome following donor embryo replacement. *J Matern Fetal Med* 1997;6:237–240
 12. Jones HW: The impact of in vitro fertilization on the practice of gynecology and obstetrics. *Int J Fertil* 1986;31:99–111
 13. Ontario Ministry of Health: Medical Policy Bulletin. Kingston, Ontario Ministry of Health, 1995
 14. SART (Society for Assisted Reproductive Technology and the American Society for Reproductive Medicine): Assisted reproductive technology in the United States and Canada: 1995 results generated from the American Society for Reproductive Medicine/Society for Assisted Reproductive Technology Registry. *Fertil Steril* 1998;69:389–396
 15. Edelmann RJ, Connolly KJ, Bartlett H: Coping strategies and psychological adjustment of couples presenting for IVF. *J Psychosom Res* 1994;38:355–364
 16. Daniels KR: Psychosocial factors for couples awaiting in vitro fertilization. *Soc Work Health Care* 1989;14:81–98
 17. Civic Hospital. Goal Program, Ottawa Civic Hospital, Ottawa: <http://www.cyberplus.ca/~conceive/>, 1996
 18. London Health Sciences Centre. Infertility Treatment: Services and Programs. London; University of Western Ontario, 1996
 19. Yuzpe AA, Brown SE, Casper RF, Nisker JA, Graves G: Rates and outcomes of pregnancies achieved in the first 4 years of an in-vitro fertilization program. *Can Med Assoc J* 1989;140:167–172
 20. Bergh C, Josefsson B, Nilsson L, Hamberger L: The success rate in Swedish in-vitro fertilization unit: A cohort study. *Acta Obstet Gynecol Scand* 1995;74:446–450
 21. Tan S-L, Doyle P, Maconochie N, *et al.*: Pregnancy and birth rates of live infants after in vitro fertilization in women with and without previous in vitro fertilization pregnancies: A study of eight thousand cycles at one center. *Am J Obstet Gynecol* 1994;170:34–40
 22. ASRM. The Society For Assisted Reproductive Technology and the American Society for Reproductive Medicine release 1993 ART success rates. Washington, DC: <http://www.asrm.com/>, 1995
 23. Taylor PJ, Mahadevan M, Leader A, Wiseman D, Lorscheider F, Servis S: Initial experience with in vitro fertilization and embryo transfer at the University of Calgary Foothills Hospital. *J In Vitro Fertil Embryo Transfer* 1985;2:112–113