Journal of Assisted Reproduction and Genetics, Vol. 18, No. 2, 2001

CLINICAL ASSISTED REPRODUCTION

IVF Births and Pregnancies: An Exploration of Two Methods of Assessment Using Life-Table Analysis

R. DEONANDAN,[1,4] M. K. CAMPBELL,[1] T. OSTBYE,[2] I. TUMMON,[3] and J. ROBERTSON[1]

Submitted: August 16, 2000
Accepted: August 16, 2000

Purpose: *Our purpose was to explore two methods of expressing the performance of IVF programs.*
Methods: *Using life-table methods, hazard and cure rates and a "monthly fecundability rate" were calculated for an Ontario IVF clinic. The rates were evaluated for their meaningfulness as indicators of the clinic's performance.*
Results and Conclusions: *While the hazard rate describes monthly fertility among those who will eventually become pregnant, the fecundability rate describes fertility for all patients who enter the program, making it the more appropriate index for program comparisons. However, from a prospective patient's perspective, both methods are valid indices for summarizing a program's performance.*

KEY WORDS: IVF; cure rate; fecundability rate; hazard rate; life-table analysis; survival analysis.

See also Editorial on page 71

[1] Department of Epidemiology & Biostatistics, University of Western Ontario, London, Ontario, Canada.
[2] Department of Obstetrics & Gynaecology, London Health Sciences Centre, London, Ontario, Canada.
[3] Department of Family Medicine, Duke University, Chapel Hill, North Carolina 27706.
[4] To whom correspondence should be addressed at 124 Coleman Avenue, Toronto, Ontario, Canada M4C 1R1. Fax: (416) 946-2291. e-mail: ray@deonandan.com.

INTRODUCTION

The life-table method is intended to provide a statistical approach to computing death rates in a study population subjected to varying lengths of follow-up. It allows researchers to measure the efficacy of a therapeutic intervention. Because it does not require all patients to enter the study simultaneously, or to be followed up for the same period of time, it is an attractive analytical tool to be applied to a host of clinical problems, including infertility.

An assumption underpinning the technique's statistical validity is that patients who drop out have the same adverse outcome rate as those who are not lost to follow-up. This is usually a valid assumption since the analysis is most often applied to the survival of populations exposed to or recovering from a single intervention, such as a surgical technique. Constant monitored care subsequent to that intervention often has little influence on the patients' survival rate.

In the realm of infertility, pregnancy and live birth, rather than death, are the outcomes of interest. Unlike the case of survival after a surgical intervention, one cannot assume that infertility patients exhibit the same chances of the outcome when absent from clinical observation and treatment. It may well be true that the women who are lost to follow-up are a low-prognosis subgroup within the treatment population (1). Indeed, patients may leave the program precisely because they are responding poorly to treatment. When such individuals are lost to follow-up, survival analysis will overestimate therapeutic efficacy. Conversely, those few patients who become pregnant spontaneously, without benefit of an in vitro fertilization (IVF) intervention, affect the analysis by possibly underestimating therapeutic success. Furthermore, IVF does not represent a single intervention but, rather, a continuous treatment involving, among other things, oocyte pickup, preparation of the zygote, implantation of the embryo, and continued psychological counseling. Some assumptions regarding life-table analysis are therefore violated, and the appropriateness of the analysis, in the absence of suitable modification, is therefore in question. The limitations of the application of life-table analysis to infertility studies have been widely commented upon (2). Modifications to the life-table method have been implemented in such research areas as clinical drug trials (3), yet there has been a dearth of such initiatives in infertility.

Despite its failings, life-table analysis continues to be applied to infertility data. This is partly because the method allows for the easy computation of two useful statistics: cumulative probabilities of conception and birth. Such rates are considered by many researchers to be a meaningful measure of program success (4–6). To this end, the life-table method is used because it takes into account the experience of the entire cohort by using all the treatment cycles.

Stolwijk *et al.* (7) correctly pointed out that two assumptions may be made regarding the group lost to follow-up, resulting in two possible outcomes. The first assumption is that the group which is lost to follow-up has the same chance of experiencing the outcome event, that is, becoming pregnant or giving birth, as the monitored group. This first assumption is implicit in the life-table analysis and, as discussed above, is inherently flawed because of the nature of the condition being treated. The second assumption is that the group which has been lost to follow-up has no chance of conceiving or of becoming pregnant. This, too, is a flawed argument since infertility is often not an absolute. Perhaps, however, the latter assumption is more defensible than the first; many cases of infertility, especially those requiring donor gametes, are so extreme that the chances of spontaneous conception, in the absence of technological assistance, are remote.

For the above reasons, it would be useful to have at hand a simple way of adjusting pregnancy rates for variability in follow-up. If traditional cumulative rates are misleading, then further indices or parameters may be extracted from the life-table process to describe the IVF scenario better. Cramer *et al.* (8) suggested a "monthly fecundability rate," f, as such a measure. Such a rate is the proportion of patients who can be expected to become pregnant or give birth, depending upon one's chosen outcome of interest, in a given month. It is defined as

$$f = \frac{\sum \text{conceptions}}{\sum \text{person-months of follow-up}}$$

and computed as part of the life-table process.

Guzick and Rock (9) lauded this rate as providing an attractive alternative to gross pregnancy rates since, in their opinion, it incorporates a simple adjustment for variation in patient follow-up. They point out, however, that f, when extrapolated, only roughly approximates the actual cumulative pregnancy curve based upon observed data. For general studies in infertility, they instead suggest the application of a parametric method which utilizes a hazard rate, λ, obtained by numerically integrating the life-table's cumulative probability curve:

$$P(t) = c[1 - e^{-\lambda t}]$$

where c is the "cure rate," or the proportion of patients who can be expected to be fertile after the application of the procedure being investigated, and $P(t)$ is the probability of birth or pregnancy as a function of time.

Both suggestions are meant to address the limitations of life-table analysis in infertility research. As a means of comparing the performances of separate programs, the method of Guzick and Rock was used on simulated data by Doody (1), while both methods were utilized by Olive and Martin (10) to explore laparoscopic treatments of endometriosis.

This study investigates the ability of both methods to describe adequately the performance of a single IVF program.

METHODS

Kaplan–Meier life-table analyses were performed on data from the IVF clinic of the London Health Sciences center. Both pregnancy and birth outcomes for 2474 patients were investigated as the end points of treatment, with the date of initial assessment as each patient's starting point.

The requirements for a life-table analysis include a well-defined starting point for each subject (8). Such definition is absent from infertility statistics. Should the starting point commence upon the patient's first assessment, even though she was likely under some kind of prereferral clinical treatment? Or, as suggested by Lenton *et al.* (11), should such an interval reflect the length of infertility reported by the patient? The latter is clearly flawed since infertility is not diagnosed until a patient actually tries to become pregnant; the underlying problem may have existed from birth or puberty.

Another obvious requirement is a well-defined end point that must be expressed as a binary outcome: "pregnant" versus "not pregnant" or "live birth" versus "no live birth." The vast majority of researchers utilizing survival analysis, including most published rates used by clinics for advertising purposes, considers a documented conception to be the valid end point. Other modified possibilities include a pregnancy beyond 20 weeks or a live birth that will remain alive for at least 28 days.

For this study, the end points of conception and live birth were deemed appropriate, largely for the convenience of data extraction. The limiting factor here is the inability of patient records to reflect a sufficient spectrum of treatment chronology: Did the live births survive for a meaningful duration? and Does initial assessment truly reflect the point of commencement of treatment? Doody (1) presents a compelling argument for a more detailed data collection system for such analytical purposes.

In addition to the standard cumulative probabilities of conception and birth, the monthly fecundability rate and hazard and cure rates were computed for evaluation and discussion. The latter two parameters were obtained by applying the Leverberg–Marquardt method of maximum-likelihood estimation as suggested by Olive and Martin (10).

RESULTS

The two parameters calculated according to the method of Guzick and Rock, and the fecundity index suggested by Cramer *et al.* (8), are summarized in Table I for both pregnancy data and birth data, along with their respective 95% confidence limits.

Figures 1 and 2 show the shapes of the curves of cumulative probabilities predicted by the Guzick and Rock method compared with the actual curves computed via life-table analysis. To test the point-by-point difference between each pair of curves, t tests were performed, resulting in a failure to detect a significant difference between the actual and the fitted points for both pregnancy ($P = 0.650$) and live birth ($P = 0.565$) outcomes.

DISCUSSION

The interpretation of the meanings of the computed parameters is an important, though somewhat subjective task. Cramer *et al.* (8) suggested that the f values of 0.043 for pregnancies and 0.025 for live births should be "readily understandable" for both clinician and patient: a given patient has a 4.3% chance of conceiving, and a 2.5% chance of giving birth, in a given month of treatment.

Any interpretation of Guzick's and Rock's parameters is more problematic. The original authors stated that the cure rate, c, referred to the proportion of patients in a given month who had had their *potential* for fertility restored. While this makes sense for other treatments of infertility, it is not as easily interpretable for IVF programs. Perhaps for IVF it is more meaningful to consider the cure rate as representing a proportion of the infertile who will eventually experience a positive outcome, regardless of how long it takes. In this regard, the cure rate can be considered to be an important measure in the eyes of a potential patient.

The hazard rate, λ, is more traditionally interpreted as the "instantaneous" probability of pregnancy or live birth, a measure congruent with f. Interestingly, for both outcomes, λ and f were significantly different, though both purport to measure the same trend. This disparity can be explained by the fact that f describes the fecundability rate for all patients entering the program, while λ describes the monthly fecundability rate among the "cured," or among those who will eventually succeed. By this reasoning, f should be considered the more valid summarizing parameter.

Table I. Parameters Computed from Life-Table Analysis with 95% Confidence Limits[a]

	Cure rate, c	Hazard rate, λ	Monthly fecundability rate, f
Pregnancy as outcome	0.586 (0.561, 0.611)	0.106 (0.072, 0.140)	0.043 (0.039, 0.047)
Live birth as outcome	0.454 (0.439, 0.468)	0.094 (0.073, 0.115)	0.025 (0.022, 0.028)

[a] Cure and hazard rates were estimated using the Leverberg–Marquardt maximum-likelihood method.

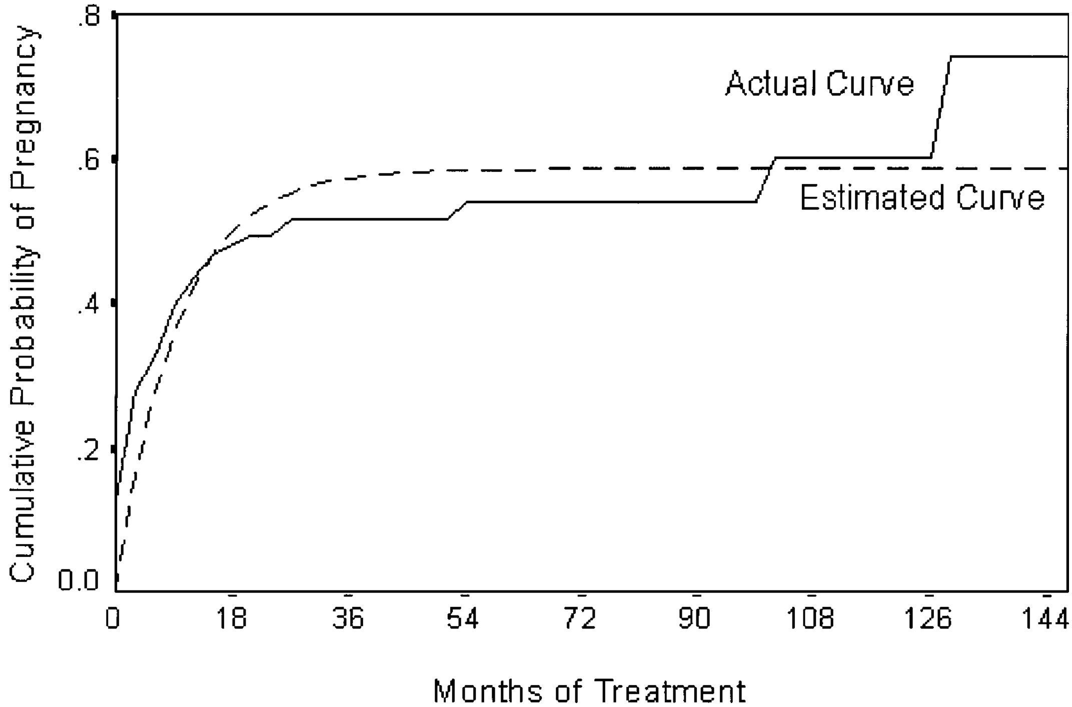

Fig. 1. Actual cumulative pregnancy curve alongside curve estimated by two-parameter method of Guzick and Rock (1981). Curve reflects data from 2474 patients.

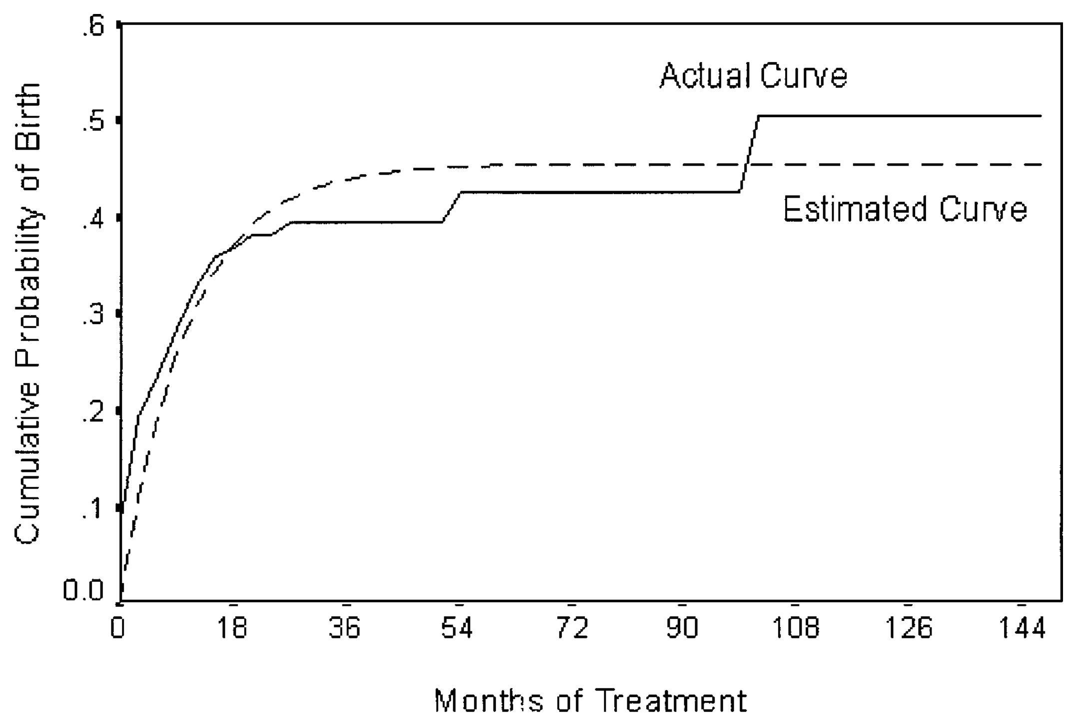

Fig. 2. Actual cumulative live birth curve alongside curve estimated by two-parameter method of Guzick and Rock (1981). Curve reflects data from 2474 patients.

Guzick's and Rock's contention that their model better estimates the actual cumulative probability curve is tested qualitatively in Figs. 1 and 2, and quantitatively via *t* tests. Clearly, the two-parameter model is an excellent estimate of the life-table curve, but this does not necessarily support the authors' claim that their parameters are the most appropriate indices for comparing performance between programs.

CONCLUSION

Both methods explored in this study provide powerful tools for the comparison of IVF programs. For administrative purposes, Cramer and co-workers' (8) monthly fecundability rate seems to be the most valid index because of its inclusion of patients who will eventually fail. For a potential consumer investigating a single program, however, Guzick's and Rock's cure and hazard rates provide meaningful descriptors of an individual's chances of success.

ACKNOWLEDGMENTS

The authors wish to thank Dr. Francis Tekpetey and the staff of the London Health Sciences Centre's IVF Clinic for their assistance in obtaining data. Thanks are also given to Dr. Michael Eliasziw for reviewing the statistical validity of the manuscript.

REFERENCES

1. Doody MC: Drop-out behaviour and fertility table analysis of pregnancy rates. Hum Reprod 1993;8:886–889
2. Wilcox LS, Peterson HB, *et al.*: Defining and interpreting pregnancy success rates for in vitro fertilization. Fertil Steril 1993;60(1):18–25
3. Huang X, Chen S, *et al.*: Piecewise exponential survival trees with time-dependent covariates. Biometrics 1998;54(4):1420–1433
4. 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
5. Alsalili M, Yuzpe AA, Tummon IS, *et al.*: Confounding variables affecting in vitro fertilization success: A decade of experience. J Assist Reprod Genet 1995;12:88–92
6. 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 pergnancies: A study of eight thousand cycles at one center. Am J Obstet Gynecol 1994;170:34–40
7. Stolwijk A, Hamilton C, Hollanders J, Bastiaans L, Zielhuis G: A more realistic approach to the cumulative pregnancy rate after in-vitro fertilization. Hum Reprod 1996;11:660–663
8. Cramer DW, Walker AM, Schiff I: Statistical methods in evaluating the outcome of infertility therapy. Fertil Steril 1979;32:80–86
9. Guzick DS, Rock JA: Estimation of a model of cumulative pregnancy following infertility therapy. Am J Obstet Gynecol 1981;140:573–578
10. Olive DL, Martin DC: Treatment of endometriosis—associated infertility with CO2 laser laparascopy: The use of one- and two-parameter eponential models. Fertil Steril 1987;48:18–23.
11. Lenton E, Weston G, Cooke I: Long-term followup of the apparently normal couple with a complaint of infertility. Fertil Steril 1977;28:913