Human Reproduction Vol.17, No.1 pp. 103–106, 2002

Multivariate Markov chain analysis of the probability of pregnancy in infertile couples undergoing assisted reproduction

Joseph McDonnell[1,3], Angelique J.Goverde[2], Frans F.H.Rutten[1] and Jan P.W.Vermeiden[2]

[1]Institute for Medical Technology Assessment, Erasmus University, Rotterdam and [2]Division of Reproductive Endocrinology and Fertility, Institute for Endocrinology, Reproduction and Metabolism, Vrije Universiteit Medical Centre, Amsterdam, The Netherlands

[3]To whom correspondence should be addressed at: Institute for Medical Technology Assessment, Erasmus University, PO Box 1730, 3000 DR Rotterdam, The Netherlands. E-mail: mcdonnell@bmg.eur.nl

BACKGROUND: Estimating the probability of pregnancy leading to delivery and the influence of clinical factors on that probability is of fundamental importance in the treatment counselling of infertile couples. A variety of statistical techniques have been used to analyse fertility data, many borrowed from survival analysis. METHODS AND RESULTS: We propose an alternative method of analysis which is based on a discrete time Markov chain approach, with states 'pregnancy (leading to a delivery)', 'not pregnant', and 'censored' and in which the transition probabilities are dependent both on the clinical characteristics of the patient and the treatment given. CONCLUSIONS: We believe that the method of analysis presented here may be preferable to standard analyses in that it better reflects the clinical situation, it is a truly discrete time analysis applied to a discrete time situation, it explicitly models the censoring process (a process which in itself provides information of interest to the physician) and can be readily extended to a variety of clinical situations.

Key words: Markov chain/pregnancy data/statistical analysis

Introduction

Estimating the probability of pregnancy leading to delivery and the influence of clinical factors on that probability is of fundamental importance in the treatment counselling of infertile couples. A variety of statistical techniques have been used to analyse fertility data, many of which have been borrowed from survival analysis [life table analysis (Land *et al.*, 1997), parametric survival analysis (Duleba *et al.*, 1992), Cox regression (Collins *et al.*, 1995; Eimers *et al.*, 1994)] or which have been used to model cyclical aspects of conception [logistic regression (Burns *et al.*, 1994; Roseboom *et al.*, 1995)]. Indeed, assisted reproduction data resemble survival data in that the primary outcome is bivariate in nature, consisting of a binary component (delivery or not) together with a component reflecting an aspect of time (number of cycles). Moreover, this outcome variable is often subject to censoring as couples may drop out before pregnancy is achieved. Experience shows that the degree of censoring is often substantial (Land *et al.*, 1997). However, the statistical methods typically used do not explicitly model the censoring process, a process which itself may contain information useful to a physician, nor do they readily allow incorporation of the results in cost- and cost-effectiveness analyses. In this study, we propose a method of statistical analysis based on a discrete-time Markov chain in which the transition probabilities are estimated by means of a multivariate regression function on relevant patient characteristics and treatment. We illustrate the use of this procedure with the analysis of data from a study conducted in the Netherlands into the cost-effectiveness of intrauterine insemination and IVF.

Materials and methods

Clinical trial

Couples with idiopathic subfertility of at least 3 years duration (n = 181) or with male subfertility of at least 1 year duration (n = 77) who had given their informed consent for participation in the trial were randomly allocated into one of three treatment programs of either intrauterine insemination in a spontaneous cycle (IUI–), IUI after mild ovarian hyperstimulation (IUI+), or IVF. Each treatment programme consisted of a maximum of six treatment cycles without cross-over. Clinical information, including the age of the female patient, the duration, type and indication of infertility, was recorded for each couple. Further details on the inclusion and exclusion criteria, randomization process and treatment protocols and are described by Goverde (Goverde *et al.*, 2000).

Multivariate Markov chain analysis

The statistical analysis, based on a Markov chain approach, was carried out to address the question as to whether the three types of

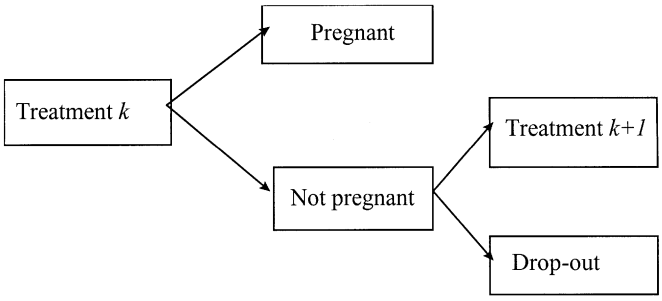

Figure 1. Treatment and outcome schema

treatment are equally effective in leading to delivery. The Markov chain analysis allows explicit modelling of the process by which couples become censored before achieving delivery. To do so, specific 'states' were identified and the probabilities of transition from one state to the other were calculated. In this analysis, three states were defined: 'pregnant' (or more exactly, 'pregnant as a result of treatment, and leading to a delivery', 'not pregnant' (but still in the treatment programme) and 'censored' [not pregnant but no longer in the programme (drop-out)]; spontaneous pregnancies are considered to be censored. 'Pregnant' and 'censored' are absorbing states, i.e. movement to another state from within these states is assumed to be impossible. Transition probabilities are assumed to be possibly dependent on certain clinical characteristics of the couple involved, for example, the age of the female, the indication of infertility (idiopathic or male subfertility) and the particular treatment given to that couple.

After entry into the study, a couple undergoes at least one round of treatment. As a result of that treatment, the female patient may become pregnant, eventually delivering a baby. If so, the couple leaves the study. If not, another round of treatment is offered. The couple may decide not to accept the offer of further treatment and therefore leaves the study; such couples are considered to be censored at the last attempt, as are couples who achieved pregnancy by natural means between treatment rounds (spontaneous pregnancy). A schema for the course of treatment and outcome is given in Figure 1.

Regression analysis

The transition probabilities are, for each couple and for each cycle, assumed to be constant over cycles (i.e. small ageing effects are ignored) and to be logistic in form. More explicitly, we assume that the probability (per cycle) of pregnancy is:

$$P[\text{pregnancy}] = P_D = \exp(\alpha_1 + \beta_1 \cdot \mathbf{x})/(1 + \exp(\alpha_1 + \beta_1 \cdot \mathbf{x}))$$

and that the probability of censoring is:

$$P[\text{censored}] = P_C = \exp(\alpha_2 + \beta_2 \cdot \mathbf{x})/(1 + \exp(\alpha_2 + \beta_2 \cdot \mathbf{x}))$$

where $\mathbf{x}$ is the vector of patient characteristics (including treatment) and $\{\alpha_1,\beta_1,\alpha_2,\beta_2\}$ are unknown coefficients.

Estimation of the coefficients is done by the method of maximum likelihood.

Construction of the likelihood

Each couple contributes to the likelihood function according to their progress (series of states). As the time process is discrete, the contribution to the likelihood made by an individual couple is simply the probability of the observed series of states. The likelihood contributions are (for each of the possible end-states):

1. Pregnant: the couple has undergone n attempts, the first $(n - 1)$ being unsuccessful, the last successful:

$$(1-P_D)(1-P_C)^{n-1}(1-P_D)^{n-2}P_D = ((1-P_D)(1-P_C))^{n-1}P_D$$

The construction of this term is as follows: if $n \geqslant 2$, then the couple have had n treatments, the first of which was unsuccessful [with probability $(1 - P_D)$], followed by $(n - 1)$ treatments, all uncensored [with probability $(1 - P_C)^{n-1}$], and all being unsuccessful except the last [with probability $(1 - P_D)^{n-2}P_D$]. If $n = 1$, the first attempt was successful and the contribution to the likelihood is simply P_D, to which the above term reduces when $n = 1$.

2. Not pregnant but did not complete all six treatments (censored):

$$(1 - P_D)((1 - P_C)(1 - P_D))^{n-1}P_C = (1 - P_D)^n(1 - P_C)^{n-1}P_C$$

The likelihood of such a sequence is as follows: the couple underwent n $(1 \leqslant n < 6)$ treatments, the first of which was unsuccessful [with probability $(1-P_D)$], followed by $(n-1)$ treatments, all uncensored [with probability $(1 - P_C)^{n-1}$] and all unsuccessful [with probability $(1 - P_D)^{n-1}$], followed by withdrawal from the study (with probability P_C).

3. Not pregnant but completed all six treatments ($n = 6$):

$$(1-P_D)((1-P_C)(1-P_D))^{n-1} = (1-P_D)^n(1-P_C)^{n-1} = (1-P_D)^6(1-P_C)^5$$

The likelihood of this sequence is derived in a fashion similar to the previous one.

Each couple contributes exactly one of the above terms to the likelihood. The likelihood of the observed data is simply the product of the individual contributions over all couples; this likelihood, or more exactly its log, was maximized to find the parameter estimates (Muenz and Rubinstein, 1985).

Each explanatory variable may make a significant contribution to both regression functions, to only one or to neither.

The statistical significance of each variable in a particular regression function was examined by comparing twice the difference in the log-likelihood computed from two models, one containing the variable in that regression function, the other with the variable removed from the function, with critical values of a ‰Σ0000Tχ^2 distribution with the appropriate degrees of freedom, as determined by the dimension of the variable.

Software

All computations were performed using the LE (Maximum Likelihood Estimation) module of BMDP (Dixon, 1990).

Results

The clinical results are presented only briefly here. For a more detailed presentation, see Goverde *et al.* (Goverde *et al.*, 2000).

Between February 1992 and September 1995, 86 couples were assigned to the IUI– group, 85 to the IUI+ group and 87 to the IVF group. Ten couples withdrew from the study before the initial treatment due to a spontaneous pregnancy, illness or a change of mind. Subsequently, an additional 64 (24.8% of the total) couples (13 IUI–, 14 IUI+, 37 IVF) dropped out of the study before achieving pregnancy. Treatment resulted in 89 pregnancies (25 IUI–, 31 IUI+, 33 IVF) and 107 babies (26, 40, 41 respectively).

Under the Markov chain analysis, no significant difference in the chance of pregnancy was found between the two IUI groups ($P > 0.5$); these groups were combined in subsequent analyses. The age of the female patient had a strong significant effect on the chance of pregnancy (per cycle) with older female patients being much less likely to achieve delivery than younger patients ($P < 0.01$). On a per-cycle basis, both IVF patients and IUI patients had a statistically similar chance of achieving

Table I. Regression coefficients of the final model under the Markov approach and Cox regression

		Coefficient	
		Markov chain	Cox regression
Pregnancy	Constant	–0.406	
	Age	–0.06163[b]	–0.05218[b]
	IVF	0.440[c]	0.329[d]
	Indication	–0.092[d]	–0.094[d]
Censoring	Constant	–3.013	
	IVF	1.398[a]	

[a]$P \leqslant 0.01$.
[b]$0.01 < P \leqslant 0.05$.
[c]$0.05 < P \leqslant 0.10$.
[d]$P \geqslant 0.10$.

a delivery, but IVF patients were much more likely to withdraw from treatment than IUI patients ($P < 0.001$). No other factors were found to have a significant effect on either the probability of conception per cycle or of drop-out.

For the final model, we decided to include the statistically non-significant variable *indication*. To compare the results of the Markov chain analysis with those of alternative methods, we ran the same model using Cox regression. The results for both models are given in Table I.

The models accorded the variables the same rank order of importance, although the *P*-values derived under the Markov model were smaller than under the Cox model. In particular, the treatment variable attained a very low degree of significance under the Cox model (see Table I, *P*-value Markov = 0.085, Cox = 0.15).

Discussion

In this paper, we present an alternative statistical approach to the analysis of pregnancy and delivery data. This model has a number of features potentially attractive to analysts. First, the Markov model is an inherently discrete model modelling an inherently discrete situation; in contrast, parametric and Cox regression models, as implemented in most statistical packages, are continuous time models which, in the context of assisted reproduction, are applied to a discrete situation. Secondly, we believe that structure of the statistical model should be based on clinical considerations and that the structure of this model corresponds closely to the physician's perception of the assisted reproduction process. Thirdly, the Markov process explicitly models the censoring process. Such a model contains information useful to the physician, who can gain insight into factors affecting withdrawal. Alternative techniques, such as Cox regression, do allow the modelling of the censoring process, but at some effort to the analyst. Modelling both the pregnancy and censoring processes facilitates cost-effectiveness analysis. Cost-effectiveness analyses involve, by definition, computing the ratios of costs and effects, both of which are subject to censoring in the assisted reproductive technique situation. Estimation of (cost-effectiveness) ratios and their variances under censoring are problematic, to say the least. By modelling both the probability of conception and censoring, it is possible to evaluate the cost-effectiveness of various treatments regimens and to determine the most cost-effective course of treatment for a couple.

The analysis presented here can also easily be adapted to other protocols and situations, which, in general, can be encompassed in other statistical models only with some difficulty. Although not utilized in the present analyses, the Markov model presented here can incorporate a number of other features: (i) Cycle- (time-) dependent co-variates such as cycle specific data (e.g. sperm count) and 'phase' data applicable to groups of cycles (e.g. a low probability of withdrawal for the first few IVF cycles but a higher probability of withdrawal for later cycles) can be included; (ii) non-proportional and other hazard structures and other forms of the cycle probabilities can be used (for example, extreme value); (iii) crossovers in treatments or more complex treatment regimens can be accommodated. For example, a trial protocol may stipulate that each participant first undergo one IVF attempt, which would yield diagnostic information on the female patient even if it is unsuccessful. Patients may then be assigned to a particular treatment group on the basis of this information. This analysis can encompass such treatment combinations simply by building the likelihood in the appropriate manner; (iv) in principle, it is simple to incorporate non-linear functions of variables in the regression functions. For example, one might expect that the probability of conception may increase as sperm count increases but that a 'ceiling effect' might appear in that counts higher than a certain value may not result in a significant increase in the probability of conception. This can be modelled by including sperm count (S), say, in the form $(1 - 1/S)$ where λ (>0) is a scaling parameter; and (v) the above approach is relevant to the clinical trial situation where the time between treatment cycles is relatively short and the chance of spontaneous pregnancy is small. In long-term follow-up studies, the probability of spontaneous pregnancy may be significant. The likelihood approach can be easily adapted to this situation. For example, suppose we wish to study factors influencing the chance of spontaneous pregnancy in couples not undergoing assisted reproductive techniques. We again define P[spontaneous pregnancy] = P_D (where P_D is a function of **x**, a vector now relating purely to patient characteristics) and P[censored] = P_C. If a given couple have been observed for *n* cycles and finally achieve pregnancy, the likelihood contribution is $((1 - P_D)(1 - P_C))^{n-1}P_D$, which is structurally identical to that in the CRT situation. The likelihood function can be easily extended to include both treatment and spontaneous cycles.

In conclusion, we believe that the Markov chain model offers a useful alternative method of analysis for conception data. The model structure corresponds closely to the clinical situation and can be readily adapted to other clinical situations. In particular, it explicitly models the censoring process, providing useful information to the physician and facilitating cost-effectiveness considerations.

References

Burns, W.N., Witz, C.A., Klein, N.A. *et al.* (1994) Serum progesterone concentrations on the day after human chorionic gonadotropin administration

and progesterone/oocyte ratios predict *in vitro* fertilization/embryo transfer outcome. *J. Assist. Reprod. Genet.*, **11**, 17–23.

Collins, J.A., Burrows, E.A. and Willan, A.R. (1995) The prognosis for live birth among untreated infertile couples. *Fertil. Steril.,* **64**, 22–28.

Dixon, W.J. (1990) *BMDP Statistical Software Manual*. University of California Press, California.

Duleba, A.J., Rowe, T.C., Ma, P. *et al.* (1992) Prognostic factors is assessment and management of male infertility. *Hum. Reprod.,* **7**, 1388–1393.

Eimers, J.M., te Velde, E.R., Gerritse, R. *et al.* (1994) The prediction of the chance to conceive in subfertile couples. *Fertil. Steril.,* **61**, 44–51.

Goverde, A.J., McDonnell, J., Vermeiden, J.P.W. *et al.* (2000) Intrauterine insemination or in-vitro fertilisation in idiopathic subfertility and male subfertility: a randomised trial and cost-effectiveness analysis. *Lancet,* **355**, 13–18.

Land, J.A., Courtar, D.A. and Evers, J.L.H. (1997) Patient dropout in an assisted reproductive technology program: implications for pregnancy rates. *Fertil. Steril.,* **68**, 278–281.

Muenz, L.R. and Rubinstein, L.V. (1985) Markov models for covariate dependence of binary sequences. *Biometrics*, **41**, 91–101.

Roseboom, T.J., Vermeiden, J.P.W., Schoute, E. *et al.* (1995) The probability of pregnancy after embryo transfer is affected by the age of the patient, cause of infertility, number of embryos transferred and the average morphology score, as revealed by multiple logistic regression analysis. *Hum. Reprod.,* **10**, 3035–3041.

Received on April 27, 2001; accepted on August 31, 2001