What is the most accurate estimate of pregnancy rates in IVF dropouts?

T.E.M. Verhagen1,3, J.C.M. Dumoulin1, J.L.H. Evers1 and J.A. Land1,2

1Research Institute Growth and Development (GROW), Department of Obstetrics and Gynaecology, Maastricht University Medical Centre, PO Box 5800, 6202 AZ Maastricht, The Netherlands
2Present address: Department of Obstetrics and Gynaecology, University Medical Centre Groningen, PO Box 30001, 9700 RB Groningen, The Netherlands
3Correspondence address. Department of Obstetrics and Gynaecology, Máxima Medical Centre, De Run 4600, PO Box 7777, 5500 MB Veldhoven, The Netherlands. E-mail: t.verhagen@mmc.nl

BACKGROUND: Dropouts in IVF-programmes affect cumulative pregnancy rates (CPRs), but it is unknown what the impact of loss to follow-up is.

METHODS: Data were obtained from 588 couples starting IVF treatment (‘as treated group’). Cycle-based and real-time-based CPRs were calculated using three assumptions for dropouts: dropouts having no probability of pregnancy, dropouts having the same probability of pregnancy as those continuing treatment and dropouts stopping because of medical reasons having no chance of pregnancy and those stopping because of other reasons having the same probability of pregnancy as those continuing treatment. CPRs obtained in the ‘as treated group’ were compared with CPRs calculated using the data set including the follow-up data of the dropouts (‘completed group’).

RESULTS: In 1.7% of couples, no follow-up could be obtained. The cycle-based CPR after three IVF-cycles ranged from 63% to 71% in the ‘as treated group’ and was 65% in the ‘completed group’. The real-time-based CPR after 9 months ranged from 54% to 59% in the ‘as treated group’ and was 55% in the ‘completed group’. The PR in dropouts was 14% (95% confidence interval 8.22%).

CONCLUSIONS: In IVF programmes, outcome data of dropouts remain unknown, and CPRs should be calculated by assuming dropouts to have a PR between no probability and the same probability as those who continue treatment. Our study shows that the most accurate estimate for the PR in dropouts is 14%.

Keywords: dropout; IVF; follow-up; life table

Introduction

Accurate information on success rates is a prerequisite for adequate counselling of couples before starting their IVF treatment. There are several methods of estimating the effectiveness of treatment. The method most often used is the life table analysis based on cumulative pregnancy rates (CPRs). This method for reporting on infertility treatment outcome was introduced >30 years ago (Lamb and Cruz, 1972). Life table analysis estimates the percentage of patients who will conceive during a particular period of follow-up, and it assumes that results in patients who dropout do not differ substantially from those who continue treatment.

It has been shown in a computer simulation program that life table analysis tends to overestimate pregnancy rates (PRs) (Doody, 1993), in particular in a population with low monthly fecundity rates and low ‘cure’ rates. These couples might be expected to be more easily discouraged from continuing treatment and to have higher dropout rates (Doody, 1993). Other studies showed that making different assumptions concerning the prognosis of couples discontinuing IVF treatment affected the CPR significantly (Stolwijk et al., 1996; Olivius et al., 2002). If one assumes that the pregnancy chance of dropouts is similar to the one in those who continue treatment, the CPR will be overestimated. On the other hand, if dropouts are assumed not to have any probability of getting pregnant, the CPR will be underestimated. It has been suggested that the most realistic estimate of the true CPR might be obtained by assuming that dropouts because of medical reasons (i.e. poor prognosis and active censoring) have no chance of pregnancy and that those who stop because of other, more personal reasons (passive censoring) have the same chance of pregnancy as those who continue treatment (Stolwijk et al., 1996; Olivius et al., 2002).

Stolwijk et al. calculated the CPR after five completed IVF cycles for these different assumptions in dropouts. The CPR was 51% when dropouts were assumed to have the same probability of pregnancy as those continuing treatment, 37% when dropouts were assumed to have no probability of pregnancy at...
all and 41% when dropouts due to passive censoring were assumed to have the same probability of pregnancy as those who continued treatment whereas actively censored dropouts had no chance of pregnancy (Stolwijk et al., 1996). Olivius et al. calculated the CPR after three completed cycles of IVF using the same assumptions. The CPR was 73% assuming all dropouts had the same probability of pregnancy as those continuing treatment, 65% assuming dropouts had no chance of pregnancy and 66% assuming dropouts due to passive censoring had the same chance of pregnancy as those who continued treatment whereas actively censored dropouts had no chance of pregnancy (Olivius et al., 2002). Both studies lacked information on the true PR in the dropouts.

In a recent review, several pitfalls of life table analysis have been summarized (Daya, 2005). Several solutions were proposed by Daya for generating more realistic outcome data. One strategy to correct for the overestimation in CPR is to use the cycle-based CPR assuming that all dropouts have a 0% PR in subsequent cycles. This approach will underestimate the CPR, but this will encourage clinics to be more realistic in counselling couples about prognosis of IVF treatment. Another strategy is the construction of a real-time-based CPR. After agreement on the average length of time a couple needs to complete one cycle of IVF treatment, involving fresh and frozen embryo transfers, the CPR can be calculated for any length of time. The outcome is summarized as the CPR after a specified period of time. However, in this strategy, there is no solution for the dropouts due to active or passive censoring. This can be solved by assuming all dropouts (whatever their reason for dropout) to have a 0% PR for the remaining study period (Daya, 2005). Daya recommends the real-time-based CPR for calculating treatment outcome because it focuses on the couple as the unit of analysis and it satisfies the time scale requirement of the life table method. The problem caused by the unknown outcome in the dropouts can only be corrected for by having access to dropout follow-up data, both with and without therapy.

In the present study, dropouts were traced and their pregnancies after dropping out of the IVF programme were calculated, in order to arrive at the best estimate of the overall CPR in couples undergoing IVF. Different methods to calculate CPR were compared and we deduced the best method for reporting IVF outcome in everyday practice, in which outcome in dropouts is usually unknown.

Materials and Methods

Data from all 588 couples starting their first IVF or ICSI-treatment in our IVF clinic between June 2000 and July 2003 were collected retrospectively. Couples starting IVF for preimplantation genetic diagnosis and surgical sperm aspiration or those using donor gametes were excluded. In the Netherlands, health retrieval reimburses three IVF cycles, and therefore a maximum of three completed cycles was included in the study. A cycle was considered completed when ovum retrieval was performed. Transfers of cryopreserved embryos were considered to be part of the fresh cycles from which they originated.

Patients were observed until either a pregnancy was achieved, until they withdrew from the programme before a pregnancy was established or until three IVF cycles had been completed. A pregnancy was defined as a positive urinary pregnancy test 14 days after embryo transfer. All pregnancies achieved during the period a couple underwent IVF were considered treatment-dependent pregnancies.

Patients withdrew from the programme when further treatment was denied for medical reasons (active censoring), or when the couple abandoned further treatment because of personal reasons (passive censoring). Reasons for active censoring were poor response, poor fertilization, poor response with poor fertilization, overweight with BMI > 30 kg/m², hypertension or improved semen quality not requiring ICSI any more. Reasons for dropout due to passive censoring were relational problems, additional health problems, psychological burden, physical burden, both psychological and physical burden, and continuation of treatment elsewhere.

Patients who did not inform staff members about their decision to abandon or postpone treatment and who failed to return to follow-up appointments for > 6 months were considered lost to follow-up.

The computer-based IVF database and the patients’ medical records were used for data collection. Follow-up data of dropouts were collected between May and December 2005 and were obtained by questionnaires sent to the patients and by contacting the patients’ general practitioner to learn whether the patient had become pregnant. All pregnancies in the dropouts were considered treatment-independent pregnancies, including pregnancies after treatment in other clinics.

For the calculations, two different data sets were constructed: (i) the outcome data of the couples who underwent IVF treatment, defined as the ‘as treated group’ and (ii) the data from the treated group supplemented with the achieved follow-up data from the dropouts, defined as the ‘completed group’.

The PR and CPR were calculated for the ‘as treated group’ and the ‘completed group’ using the cycle-based and real-time-based methods, respectively, as proposed by Daya (2005). For the cycle-based CPR, a maximum of three treatment cycles was used. For the real-time-based CPR, the length of time for one completed cycle (including fresh and frozen embryo transfers) was arbitrarily set at 3 months, since in our programme 85% of couples completed one IVF cycle within this period.

The data set of the ‘as treated group’ was used to estimate the cycle-based and real-time-based PR and CPR with the assumptions usually made for dropouts (Stolwijk et al., 1996; Olivius et al., 2002): (i) assuming dropouts had no chance of pregnancy (in the remainder of the manuscript this is abbreviated as ‘dropouts = 0%’), (ii) assuming dropouts had the same chance of pregnancy as those who continued treatment (‘dropouts = same’) and (iii) assuming dropouts due to active censoring had no chance of pregnancy and those dropping out due to passive censoring had the same chance of pregnancy as those who continued treatment (‘dropouts = differentiated’).

The data set of the ‘completed group’ was used to calculate the effective cycle-based and real-time-based PR and CPR. The CPRs from the treated group with the three different assumptions for PR in the dropouts were compared with the ‘true’ CPR from the completed group.

In the Netherlands, for retrospective chart review, no ethical review board approval is required. Before starting their IVF treatment, all couples had given written informed consent for the use of their anonymized medical data for research purposes.

The data were analysed using Microsoft Excel software, 95% confidence intervals (CI) were calculated according to Wilson’s method (Wilson, 1929).

Results

Between June 2000 and July 2003, 588 couples started their first cycle in our IVF centre. Together they completed 1024
cycles. The characteristics for the total group, the couples completing IVF treatment and those dropping out are listed in Table I. There were no significant differences in composition between the two groups of couples.

Of the 588 couples starting IVF, 480 (82%) completed treatment (i.e. became pregnant or completed three treatment cycles) and 108 (18%) dropped out. Before the first cycle was completed, 26 couples (24% of all dropouts) had stopped: 20 because of poor response (active censoring) and 6 for personal reasons (passive censoring). After the first cycle, another 47 couples (44% of all dropouts) withdrew from the programme (21 actively and 26 passively censored). After the second cycle, 35 couples dropped out (32%) (10 actively censored and 25 passively censored). Reasons for active censoring were poor response (n = 29), poor fertilization (n = 10), poor response with poor fertilization (n = 6), overweight with BMI > 30 kg/m² (n = 3), hypertension (n = 1) and improved semen quality not requiring ICSI any more (n = 2). Reasons for dropout due to passive censoring were relational problems (n = 6), additional health problems (n = 3), psychological burden (n = 8), physical burden (n = 4), both psychological and physical burden (n = 18), continuation of treatment elsewhere (n = 6) and unknown reasons (n = 12).

There were 368 pregnancies in the treated couples (‘as treated group’), 325 after fresh embryo transfers and 43 after frozen embryo transfers. These pregnancies resulted in 278 live births (76%), 251 after fresh transfers (77%) and 27 after frozen transfers (63%). Details are given in Table II.

In 98 of the 108 dropouts (91%), follow-up data were secured by sending questionnaires or contacting the general practitioners. Only 10 couples (1.7% of all couples who entered the programme) were lost to follow-up completely, of them three were actively censored and seven passively. There were 15 pregnancies in the 98 dropouts (15%): three patients conceived after fertility treatment at another clinic and 12 patients conceived spontaneously, resulting in 13 live births. Of the 51 actively censored couples (follow-up available in 48), 12 did eventually become pregnant (25%) and of the 57 passively censored couples (follow-up in 50), three conceived (6%).

### Calculation of CPR using different strategies

#### Cycle-based CPR

The cycle-based CPR and 95% CI for the ‘completed group’ was 65% (61.69%) (Table III). Table IV shows the cycle-based CPR using the data set from the ‘as treated group’. For these calculations, three different assumptions were made for the dropouts: ‘dropouts = 0%', ‘dropouts = same' and ‘dropouts = differentiated'. The cycle-based CPR and 95% CI were 63% (59.67%), 71% (67.74%) and 66% (62.70%), respectively.

Figure 1 shows the different cycle-based CPRs and illustrates that the CPR in the ‘completed group’ corresponds best with the CPR calculated by applying the ‘dropouts = differentiated’ assumption.

#### Real-time-based CPR

This strategy for calculating the CPR is based on the time to pregnancy. The time needed to complete one cycle including fresh and frozen embryo transfers was set at 3 months, and therefore the time needed for three cycles was 9 months. In Table V, the results for the ‘completed group’ are given. In one couple who stopped treatment, the time to pregnancy was not known. The CPR and 95% CI is 55% (51.59%) after 9 months for the ‘completed group’. In Table VI, the CPR for the ‘as treated group’ is given, using the three different assumptions concerning the dropouts (‘dropouts = 0%', ‘dropouts = same’ and ‘dropouts = differentiated'). These expected real-time CPRs and 95% CI were 54% (50.58%), 59% (55.63%) and 56% (52.60%), respectively, after 9 months.
Figure 2 shows the real-time-based CPR for the different subgroups for 9 months. The CPR in the ‘completed group’ corresponds best with the CPR calculated by applying the ‘dropouts = differentiated’ assumption.

**Discussion**

The cycle-based and real-time-based CPRs were calculated for the ‘as treated group’ and the ‘completed group’, using data from 588 couples. The ‘completed group’ included the follow-up data from dropouts. In most studies, performed a substantial number of follow-up data in dropouts is missing. In our study, follow-up was complete in 98.3% of all couples starting treatment, and in 98 of 108 couples who dropped out (91%) follow-up data could be secured. In this ‘completed group’, a 65% cycle-based CPR after three cycles and a 55% real-time-based CPR after 9 months were calculated. For the ‘as treated group’, CPR was calculated using different assumptions for the dropouts. For the cycle-based CPR, the ‘dropouts = differentiated’ assumption was found to
correspond to the cycle-based CPR in the ‘completed group’ (65% and 66%, respectively). For the real-time-based CPR, after 9 months of treatment, the CPR in the ‘completed group’ (55%) was comparable with the CPR for ‘dropouts = 0%’ (54%) and the ‘dropout = differentiated’ (56%) in the ‘as treated group’. These findings indicate that the impact of dropouts is smaller on real-time-based CPRs when compared with cycle-based CPRs.

In most of the studies presenting CPRs, different assumptions for dropouts are made. Incorporating the reason for dropout (in this study ‘dropouts = differentiated’) is assumed to give the most realistic estimation of the CPR (Stolwijk...
et al., 1996; Olivius et al., 2002), as actively censored patients are considered to have poor pregnancy chances and passively censored patients to have a better prognosis. Although our study confirmed that the strategy in which the reason for dropout was taken into account gave the most realistic CPR, the generally held assumption concerning PR in actively and passively censored patients could not be corroborated. There were 12 pregnancies in 48 actively censored couples (25%, 95% CI 14.40%) and only three pregnancies in 50 passively censored couples (6%, 95% CI 2.18%). There were three actively censored patients lost to follow-up and seven passively censored ones. We have performed a sensitivity analysis and assumed, according to a worst-case scenario, that none of the dropouts lost to follow-up did conceive. CPRs were also calculated after extrapolating the pregnancies in the known dropouts: 12 pregnancies in actively censored couples (follow-up in 48 of 51 couples) and three in passively censored couples (follow-up in 50 of 57 couples). Extrapolation of these data meant one more pregnancy in the actively censored couples. This did not affect the CPRs. We did not perform a best-case scenario since the findings in the known dropouts indicated a high pregnancy rate in the dropouts lost to follow-up to be unrealistic, but even if all seven passively censored patients would have conceived, the resulting PR of 20% in the passively censored patients would not have exceeded the PR of 25% in the actively censored group. Reasons for the PR in actively censored patients to be higher than in the passively censored group in our study might be due to (too) strict criteria for active censoring, or to passively censored couples refraining from trying to conceive after censoring. This study offers support for the contention that the differentiated method offers the best estimate of the eventual CPR, both cycle and real-time based. From the group with complete follow-up in our study, we calculated the PR in dropouts to be 14% (95% CI 8.22%).

In our study, a pregnancy was defined as a positive urinary pregnancy test 14 days after the embryo transfer. We are aware that for counselling couples cumulative ongoing pregnancy rates or live birth rates are preferred (Min et al., 1996; Olivius et al., 2002). In our previous study, in 1997, a dropout rate of 26% was found after the first cycle, and 32% of these dropouts were due to active censoring. After the second cycle, the dropout rate was 29%, of which 29% was due to active censoring (Land et al., 1997). These dropout rates are in accordance with other studies published so far, reporting rates between 22% and 40% after the first IVF cycle (Roest et al., 1998; de Vries et al., 1999; Olivius et al., 2002; Schroder et al., 2004). In the present study, the dropout rates were lower, namely 12% after the first and 12% after the second cycle, and the percentage of actively censored patients was 56 and 29, respectively. The decreasing percentage of couples dropping out from our IVF programme between the two study periods might be due to better patient counselling, surveillance of their progress and more structural support for couples between treatment cycles. This might have decreased dropping out for reasons related to the psychological and physical burden of the IVF treatment. We strive for minimizing passive censoring, since our couples have taught us the importance of actively putting a period to an end in order to be able to cope with their eventual persisting childlessness.

In conclusion, for counselling couples who start treatment and for reporting on IVF-results, CPR is to be preferred instead of PR per cycle. In countries in which the number of treatment cycles is fixed, cycle-based CPR is the best outcome parameter for reports on IVF. Overestimation or underestimation of CPR due to invalid assumptions of PR in dropouts should be minimized and CPR should be calculated by assuming dropouts to have a PR between no probability of pregnancy at all and the same probability of pregnancy as those who continue treatment. Our study shows that the most accurate estimate of the PR in dropout is 14% (95% CI 8.22%).

Acknowledgements

The authors acknowledge the nurses of the IVF unit in the university hospital Maastricht (L. Colautti, G. Kengen, C. Schroen, M.J. Steijns and C. Wolfs) whose excellent patient contacts form the basis for the present study with a nearly complete follow-up.

References

Daya S. Life table (survival) analysis to generate cumulative pregnancy rates in assisted reproduction: are we overestimating our success rates? Hum Reprod 2005;20:1135–1143.


Submitted on December 12, 2007; resubmitted on April 23, 2008; accepted on May 1, 2008