Human Reproduction, Vol.1, No.1 pp. 1–8, 2009
doi:10.1093/humrep/dep041

human reproduction

ORIGINAL ARTICLE *Infertility*

Cost-effectiveness of seven IVF strategies: results of a Markov decision-analytic model

Audrey A.A. Fiddelers[1,4,6], Carmen D. Dirksen[1], John C.M. Dumoulin[2], Aafke P.A. van Montfoort[2], Jolande A. Land[2,5], J. Marij Janssen[2], Johannes L.H. Evers[2], and Johan L. Severens[1,3]

[1]Department of Clinical Epidemiology and Medical Technology Assessment, Academic Hospital Maastricht, PO Box 5800, 6202 AZ Maastricht, The Netherlands [2]Department of Obstetrics and Gynaecology, Academic Hospital Maastricht, PO Box 5800, 6202 AZ Maastricht, The Netherlands [3]Department of Health Organisation, Policy and Economics, Maastricht University, PO Box 616, 6200 MD Maastricht, The Netherlands [4]Present address: Department of Anaesthesiology, Academic Hospital Maastricht, PO Box 5800, 6202 AZ Maastricht, The Netherlands [5]Present address: Department of Obstetrics and Gynaecology, University Medical Center Groningen, PO Box 30001, 7900 RB Groningen, The Netherlands

[6]Correspondence address. E-mail: Audrey.fiddelers@mumc.nl

BACKGROUND: A selective switch to elective single embryo transfer (eSET) in IVF has been suggested to prevent complications of fertility treatment for both mother and infants. We compared seven IVF strategies concerning their cost-effectiveness using a Markov model.

METHODS: The model was based on a three IVF-attempts time horizon and a societal perspective using real world strategies and data, comparing seven IVF strategies, concerning costs, live births and incremental cost-effectiveness ratios (ICERs).

RESULTS: In order to increase pregnancy probability, one cycle of eSET + one cycle of standard treatment policy [STP, i.e. eSET in patients <38 years of age with at least one good quality embryo and double embryo transfer (DET) in the remainder of patients] + one cycle of DET have an ICER of €16 593 compared with three cycles of eSET. Furthermore, three STP cycles have an ICER of €17 636 compared with one cycle of eSET + one cycle of STP + one cycle of DET, and three DET cycles have an ICER of €26 729 compared with three cycles STP.

CONCLUSIONS: Our study shows that in patients qualifying for IVF treatment, combining several transfer policies was not cost-effective. A choice has to be made between three cycles of eSET, STP or DET. It depends, however, on society's willingness to pay which strategy is to be preferred from a cost-effectiveness point of view.

Key words: Markov model / cost-effectiveness / single embryo transfer / IVF

Introduction

For many years in IVF programs, transfer of more than one embryo has been common practice. Data from the European IVF registry in 2003 show pregnancy rates to be 29% of which 23% are multiple pregnancies (Andersen *et al*., 2007). Multiple pregnancies are considered one of the most important complications of infertility treatment, because of the high incidence of obstetric, perinatal and neonatal complications for both mother and infants (Land and Evers, 2003). At the same time, these complications lead to high health care costs. Multiple pregnancies can be prevented by applying elective single embryo transfer (eSET) in IVF. To date, seven studies have been performed on the cost-effectiveness of eSET versus double embryo transfer (DET) (Wolner-Hanssen and Rydhstroem, 1998; De Sutter *et al*., 2002; Gerris *et al*., 2004; Lukassen *et al*., 2005; Fiddelers *et al*., 2006; Thurin Kjellberg *et al*., 2006; Polinder *et al*., 2008). According to a recently performed review (Fiddelers *et al*., 2007) on four of these studies (Gerris *et al*., 2004; Lukassen *et al*., 2005; Fiddelers *et al*., 2006; Thurin Kjellberg *et al*., 2006), DET turned out to be both more expensive and more effective than eSET under all circumstances. eSET is only preferred from a cost-effectiveness point of view when performed in good prognosis patients and with the inclusion of subsequent frozen/thawed cycles. In all other patients, it depends on what society is willing to pay for one extra live birth (i.e. resulting in at least one live born child), whether one cycle of eSET or one cycle of DET is to be preferred from a cost-effectiveness point of view.

In most western European countries, a complete IVF treatment consists of a maximum of three IVF cycles (Andersen *et al*., 2007). Nevertheless, cost-effectiveness analyses performed so far have compared one cycle of eSET with one cycle of DET. Only one study has determined the cost-effectiveness of a maximum of three cycles of eSET in good prognosis patients and DET in a Markov model (De Sutter *et al*., 2002). According to this study, costs per live birth were equal for eSET and DET. However, the authors mention that because of several assumptions they made (i.e. constant pregnancy rates for all cycles, no cancelled cycles, etc.), the model did not allow for testing a realistic situation (De Sutter *et al*., 2002). Recently, an RCT was published in which the cost-effectiveness of up to four reimbursed cycles of eSET were compared with a maximum of three reimbursed cycles of DET (Heijnen *et al*., 2007; Polinder *et al*., 2008). However, this study focussed on the combination of mild ovarian stimulation with eSET versus standard stimulation with DET. Moreover, in both studies, the full IVF treatment consisted of either eSET or DET, whereas in daily practice combined eSET/DET strategies are usually offered (Andersen *et al*., 2007).

At the University hospital of Maastricht, we performed an RCT in which 308 couples were included who started their first IVF cycle. In cases where at least two normally fertilized embryos were available, couples were randomized between eSET and DET, irrespective of female age and embryo quality. Details of the patients and study design have been published before (Van Montfoort *et al*., 2006a). In this study, only one cycle of eSET or DET was performed, followed by the standard treatment policy (STP) used in Maastricht (i.e. eSET in patients <38 years of age with at least one good quality embryo and DET in the remainder of patients). Alongside this RCT, a cost-effectiveness analysis was performed comparing one cycle of eSET versus one cycle of DET (Fiddelers *et al*., 2006). However, in order to determine the 'real-world' cost-effectiveness of embryo transfer policies, several IVF strategies based on more IVF cycles should be compared in a comprehensive cost-effectiveness analysis, thus evaluating the full IVF procedure. For this purpose, a Markov model was developed based on a maximum of three consecutive IVF cycles, to determine and compare the cost-effectiveness of different IVF strategies from a societal perspective. The strategies included: (i) eSET in all patients, (ii) eSET in good prognosis patients (patients younger than 38 years of age with at least one good quality embryo) and DET in the remainder of patients (i.e. STP) and (iii) DET in all patients. The purpose of this model is to reflect the 'real-world' situation as accurately as possible, taking into account cancelled cycles, availability of only one embryo (compulsory SET, cSET), declining pregnancy rates in subsequent cycles, frozen embryo transfers (FETs) and treatment dropouts.

Materials and Methods

Markov model

Modeling represents the real world with a series of numbers and mathematical and statistical relationships (Brennan and Akehurst, 2000). In health economic evaluations, models are used, for example, to extrapolate beyond the period observed in clinical trials (Buxton *et al*., 1997). Besides this, modeling is considered to support policy decision-making, based on best available evidence (Sculpher *et al*., 2006).

For an introduction to modeling health care interventions, the reader is referred to a series of articles (Miller and Homan, 1994; Detsky *et al*., 1997a,b; Krahn *et al*., 1997; Naglie *et al*., 1997; Naimark *et al*., 1997) and a recent handbook (Briggs *et al*., 2006). In short, Markov models assume that a patient is always in one of a finite number of (health) states. All events are represented as transitions from one state to another. A Monte Carlo evaluation of a Markov model determines the prognoses of a large number of individual patients. The time horizon of the analysis is divided into Markov cycles. Each patient begins in an initial state. During each Markov cycle, the patient may make a transition from one state to another, as dictated by the transition probabilities. After the first patient has completed the simulation, another patient begins in the initial state and a new simulation is performed. This process is repeated a very large number of times (for example, 5000 times), and each simulation generates a cost and effectiveness outcome for that theoretical patient and mean costs and effectiveness are estimated across all patients. For assessing the quality of a model, several criteria such as model structure, data used as inputs to models and model validation have been developed (Weinstein *et al*., 2003; Philips *et al*., 2006).

Model structure

The strategies used in our model were selected based on two criteria. First, a maximum of three subsequent IVF cycles was allowed, since this is the standard policy in most European countries. Second, the strategies had to represent a similar or constructive transfer policy. Strategies representing a similar transfer policy consisted of three consecutive cycles of eSET, three consecutive cycles of DET and three consecutive cycles of STP. Strategies representing a constructive transfer policy consisted of one cycle of eSET + two cycles of STP; one cycle of eSET + two cycles of DET; one cycle of eSET + one cycle of STP + one cycle of DET and one cycle of STP + two cycles of DET. In the model comparing these seven IVF strategies, the structure represented in a decision tree was the same for all strategies (Fig. 1). The square in the Markov tree indicates a decision node after which only one decision is possible. The nodes with an 'M' indicate a Markov node after which several health states are possible. The nodes represented by circles are probable events used if subsequent outcomes are possible.

The model starts with theoretically subfertile couples included for IVF or ICSI treatment. We assumed no difference occurred in costs and effects between IVF and ICSI. The x-Markov nodes in our model each have four health states: 'hormonal stimulation', 'FET', 'stopping while having no child' and a 'live birth'. All patients start in the health state 'hormonal stimulation'. After having started hormonal stimulation, the cycle may be cancelled because of hyperstimulation or poor response, whereas the remaining patients will have an ovum pick-up (OPU). After fertilization of the oocytes by 'regular' IVF or ICSI, embryo transfer will take place. If eSET is performed, one embryo is transferred in all patients. If STP is performed, one embryo is transferred in patients younger than 38 years of age with at least one good quality embryo, and two embryos are transferred in all other patients. If DET is performed, two embryos are transferred in all patients. In all strategies, patients have a probability that no transfer will be performed because of a total fertilization failure. Furthermore, in all strategies cSET is performed in case only one embryo is available. After embryo transfer, patients may become pregnant with a single child or twins, assuming the probability for a triplet or higher order pregnancy to be zero since never more than two embryos are transferred in any of the strategies. During pregnancy or birth, complications may occur which possibly results in a miscarriage or stillbirth. A pregnancy is defined as complicated if the mother had at least one hospital admission day during pregnancy that was not related to the delivery. Complications after birth for the child(ren) are defined as having at least one hospital

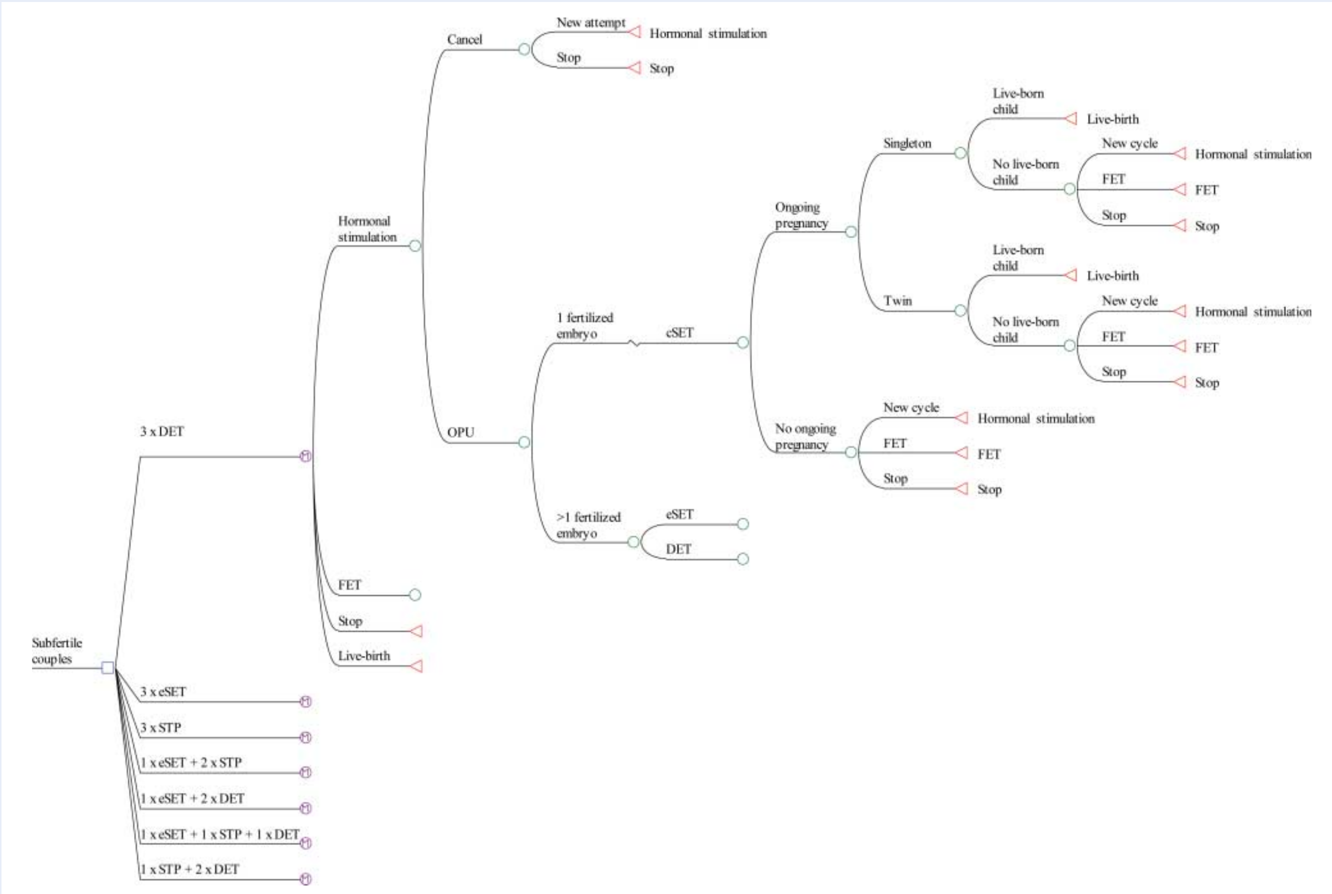

Figure 1 Markov model for cost-effectiveness of seven IVF strategies.
SET, single embryo transfer; cSET, compulsory SET; DET, double embryo transfer; eSET, elective SET; FET, frozen embryo transfer; STP, standard treatment policy; OPU, oocyte pickup.

admission day at a neonatal care unit. Patients, who did not achieve a pregnancy after embryo transfer, or had either a miscarriage or a stillborn child, may have FETs. If a patient is not pregnant after a maximum of two FETs, a new IVF cycle is started, or the IVF treatment will stop. In the Markov model a few criteria were defined for stopping IVF treatment. First, treatment ends when two subsequent IVF cycles of a certain patient are cancelled, since in this case the probability of success is assumed to be too small to allow further treatment. Second, treatment also ends if patients drop out for personal reasons. Furthermore, if a patient received the maximum of three cycles in all strategies, and a maximum of two FETs per IVF cycle, the IVF treatment is considered completed. Finally, the treatment ends if it results in a live birth.

The cycle time used in the model was defined as one IVF cycle (including fresh transfers and FETs). The time horizon was defined as a maximum of three cycles in all strategies. We have chosen this cycle time and time horizon since the calendar time duration of an IVF cycle may vary between patients.

Model input, probabilities

The probability data that were used as input for the Markov model are presented in Supplementary Table S1. The probability data were collected from several sources.

(1) Our RCT participants ($n = 308$). Data from the trial were used to determine pregnancy probabilities after the first cycle of eSET and DET (Van Montfoort *et al.*, 2006b);
(2) Non-RCT participants ($n = 222$). Non-RCT participants consisted of eligible patients who declined to participate in the RCT and eligible patients with a language barrier. Data based on simultaneously treated non-RCT-participating patients were used to determine pregnancy probabilities after receiving the first cycle of STP (Van Montfoort *et al.*, 2006b);
(3) RCT participants ($n = 308$), non-RCT participants [eligible ($n = 222$) and non-eligible patients ($n = 138$)], total $n = 668$. We used aggregated data from participating, non-participating and non-eligible patients, to diminish selection bias and to approximate reality as much as possible. Data from this population were used to determine several probabilities that were assumed to be comparable for participants, non-participants and non-eligible patients, such as the number of cancelled cycles (Supplementary Table S1);
(4) Additional patient population ($n = 153$). We used an additional patient population of all patients who had become pregnant after IVF treatment from 1995 until 2003 at the university hospital Maastricht, and who delivered at the hospital (89 singleton pregnancies and 64 twin pregnancies; in the Netherlands, 30% of singleton deliveries take place at home). Data from these patients, and their newborns were used to determine several probabilities that occurred

during pregnancy and during birth until 6 weeks post-partum (Supplementary Table S1);

(5) Literature. Data from the literature were used to determine probabilities that could not reliably be derived from the patient populations one to four described above from our university hospital owing to the small probabilities, such as the probability of birth after complications during pregnancy (Supplementary Table S1);

(6) Expert opinion. The range for pregnancy rate after eSET for use in a sensitivity analysis was determined according to expert opinion. Also, we assumed that the pregnancy rate after a FET treatment was equal after each failed eSET, irrespective of the strategy in which eSET was performed. Furthermore, FET PRs were also equal after each failed DET, after each failed eSET in STP and after each failed DET in STP, irrespective of the strategy in which the treatment was performed.

Model input, costs

The cost analysis was performed from the societal perspective and included health care costs and costs outside the health care sector (Oostenbrink *et al*., 2002, 2004; Fiddelers *et al*., 2006).

The costs were determined empirically for each couple starting IVF treatment, from 2 weeks before randomization up to 6 weeks after birth, or up to 2 weeks after the last OPU in case no pregnancy was achieved. In our model, the following costs were calculated: costs of IVF treatment (hormonal stimulation, OPU, laboratory, embryo transfer), costs of a singleton and twin pregnancy (complicated and non-complicated pregnancy), costs of delivery of a singleton and twin and costs of the period from birth until 6 weeks after birth, for the mothers as well as the children. For a detailed overview of all costs used in the model, see Supplementary Table S2. Costs were calculated by multiplying volumes of use of health care facilities [determined by the hospital information system and the patients' cost diaries (Fiddelers *et al*., 2006)] by unit prices [determined by hospital specific unit prices, micro costing calculations and guideline prices obtained from Oostenbrink *et al*. (2004)]. These prices are also listed in Supplementary Table S2. All costs were determined for the year 2003.

Model analysis

The outcome measures of the economic evaluation were the costs and effectiveness of each IVF strategy. Effectiveness was expressed as a live birth (i.e. resulting in at least one live born child). Based on costs and effectiveness of each IVF strategy included in the model, incremental cost-effectiveness ratios (ICERs) were calculated, expressing the extra costs per additional live birth. The baseline values of the probabilities and costs were incorporated into the Markov model by using the software program DATA version Pro (TreeAge software, Williamstown, MA, USA). The model was validated by experts, before it was analyzed by both first and second order Monte Carlo simulation. Furthermore, the model was debugged to check the accuracy of the equations in such a manner that the model output can be predicted.

In the first order Monte Carlo analysis, a number of individual patients were simulated. This analysis measures the variability between patients (within one sample coming from one population). The variability between the patients is caused by a random number generator. In this analysis, all parameters (probabilities and costs) are fixed. The model turned out to be stable after simulating 5000 patients.

The Markov model was also analyzed by second order Monte Carlo simulation, i.e. probabilistic sensitivity analysis to test parameter uncertainty (variability between different samples coming from one population). For this purpose, distributions were fitted for all parameters in the model. In decision theory, it is common to use a beta distribution to represent a parameter whose value is constrained between 0 and 1 (Briggs *et al*., 2006). Therefore, beta distributions were fitted for probabilities. For costs (and not for volumes of care and cost prices separately), gamma distributions were fitted. These distributions are commonly used for parameters that vary between 0 and infinite. For some probabilities and cost parameters we were not able to fit a beta or gamma distribution, since these parameters were fixed, or consisted only of a minimum and maximum variant. We therefore fitted a uniform distribution for these parameters. A uniform distribution has a low and high value, and the probability for having the low value is the same as for having the high value, and for having each value in between. The values of n (total population) and R (cases) for the beta distributions, $\alpha(\mu^2/s^2)$ and $\beta(s^2/\mu)$ for the gamma distributions and the lowest and highest values for the uniform distributions, are listed in Supplementary Tables S1 and SS2, where the sources of these distributions are also listed. For the probabilistic sensitivity analysis, 1000 iterations of 5000 patients were performed. In every iteration, for each parameter in the model, a probability or cost value was randomly picked from the specific distribution.

Cost-effectiveness acceptability curves

The choice of a particular treatment strategy depends on what society is prepared to pay for a gain in effectiveness, which is called the ceiling ratio. In other words, the probability that a treatment is cost-effective varies depending on the ceiling ratio. This can be shown in a cost-effectiveness acceptability curve (Van Hout *et al*., 1994; Fenwick *et al*., 2001). These curves were derived using the results of the probabilistic sensitivity analysis for different levels of ceiling ratios. For each strategy, the net monetary benefit was calculated by subtracting the costs from the effect multiplied by the ceiling ratio, for 1000 iterations. Per iteration, the strategy with the highest net monetary benefit is preferred and receives the value 1. The other strategies receive the value 0. Thus the probability that a strategy is most cost-effective over all the others can be determined for each particular ceiling ratio. This is repeated for several ceiling ratios, varying from €0 to €100 000 in order to construct the cost-effectiveness acceptability curves for each of the seven strategies. Based on these curves, the cost-effectiveness frontier can be determined, indicating which strategy is to be preferred above the other one (Fenwick *et al*., 2001).

Results

See Supplementary Table S3 for the results of the first order Monte Carlo simulation and of the sensitivity analyses.

Second order Monte Carlo simulation

Expected costs' and expected effectiveness per patient for each strategy are presented in Table I. The ICERs are the incremental costs per extra unit of effect, comparing different IVF strategies. The ICERs based on mean costs and mean effects show that the strategies consisting of one cycle of eSET + one cycle of STP + one cycle of DET, and one cycle of STP + two cycles of DET are ruled out as extended dominated. Furthermore, the one cycle of eSET + two cycles of DET strategy is ruled out as dominated. Comparing a one-cycle eSET + two-cycle STP strategy with a three-cycle eSET strategy results in an ICER of €7405, comparing a three-cycle STP strategy with a one-cycle eSET + two-cycle STP strategy results in an ICER of €8190 and comparing a three-cycle DET strategy with a three-cycle STP strategy results in an ICER of €17 746.

Table I **Second order Monte Carlo simulation, societal perspective mean costs, effects, cost-effectiveness and incremental costs per effect of the SET and DET strategy**

Strategy	Mean effect per couple[1] (%) Mean (95% CI)	Mean cost per couple per strategy[2] (€) Mean (95% CI)	ICER (€/effect)	
1. 3 × eSET	0.396 (0.312–0.475)	16 381 (12 544–21 861)		
2. eSET + 2 × STP	0.433 (0.364–0.499)	16 655 (12 893–21 819)	**7405**	
3. eSET + STP + DET	0.443 (0.384–0.507)	17 092 (13 198–22 736)	43 700	Dom[4] by 5
4. eSET + 2 × DET	0.457 (0.386–0.530)	17 440 (13 440–22 702)	24 857	Dom[4] by 5
5. 3 × STP	0.475 (0.414–0.537)	16 999 (13 308–22 505)	**8190**	
6. STP + 2 × DET	0.499 (0.442–0.561)	17 444 (13 681–22 546)	18 542	Ext dom[3] by 7
7. 3 × DET	0.534 (0.444–0.618)	18 046 (13 733–24 210)	17 200	**17 746**

The difference in effects and costs between first and second Monte Carlo analysis is caused by using fixed parameters for evaluating the model in first order Monte Carlo analysis, whereas oblique and uniform distributions were fitted in the second order analysis (Briggs *et al*., 2006). Definite ICERs represented in bold. SET, single embryo transfer; DET, double embryo transfer; eSET, elective SET; STP, standard treatment policy.
[1]Effect per couple: mean live birth probability for a couple starting IVF treatment.
[2]Cost per strategy includes cost of the IVF treatment, pregnancy and birth until 6 weeks later.
[3]Ext dom means extended dominated; i.e. the strategy has higher cost per pregnancy compared with the next strategy.
[4]Dom means dominated; i.e. the strategy has lower costs and a higher effect compared with the next strategy.

Cost-effectiveness acceptability curves

The cost-effectiveness acceptability curves are shown in Fig. 2. Until the ceiling ratio (expressing society's willingness to pay for a successful IVF attempt) reaches €7350, the probability that three-cycle eSET strategy is most cost-effective is the highest. When the ceiling ratio is between €7350 and €15 250, the probability that three-cycle STP strategy is most cost-effective is the highest, and when the ceiling ratio is above €15 250, the probability that three-cycle DET strategy is most cost-effective is the highest. The four strategies in which a combination of several transfer policies was combined (one cycle of eSET + two cycles of STP; one cycle of eSET + two cycles of DET; one cycle of eSET + one cycle of STP + one cycle of DET; one cycle of STP + two cycles of DET), never reached the highest probability to be cost-effective compared with the three strategies mentioned above. The cost-effectiveness acceptability frontier (not shown), followed the eSET curve until the ceiling ratio reached €7350, followed the STP curve until a ceiling ratio of €15 250 and followed the DET curve for all values above €15 250.

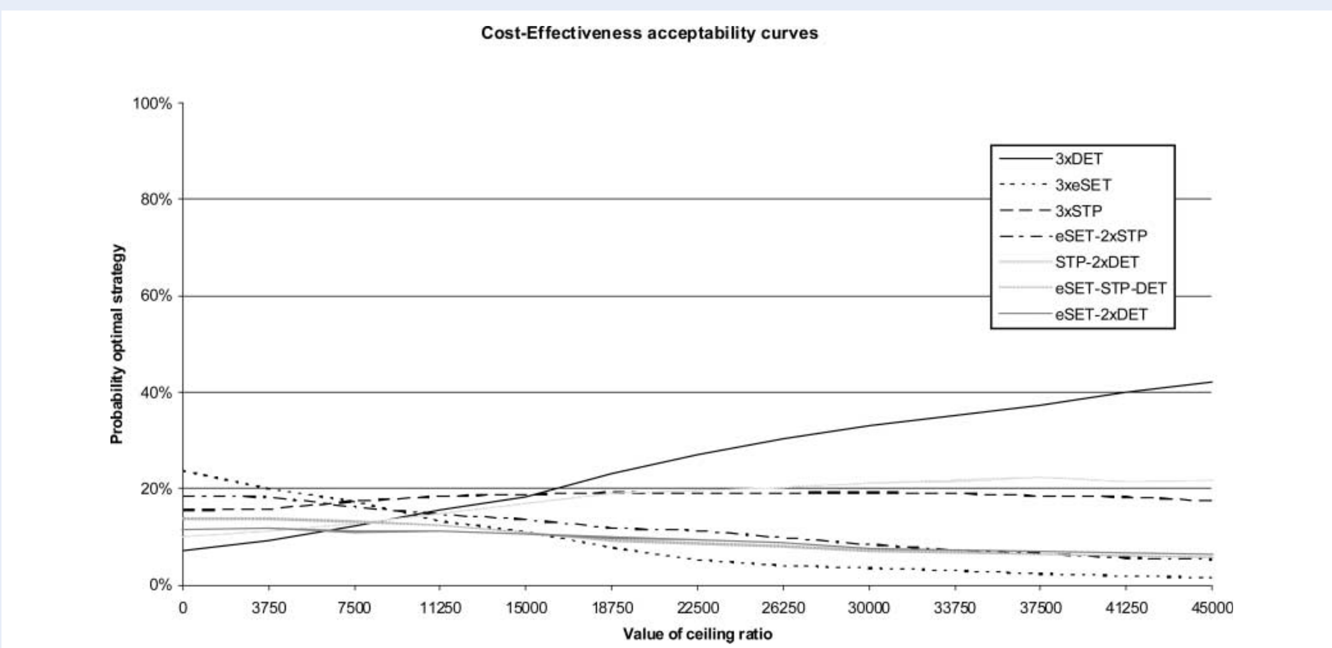

Figure 2 Cost-effectiveness acceptability curves.

Discussion

This is the first study in which a Markov model was used to determine the costs per live birth of several IVF strategies in one analysis, evaluating the full IVF procedure. Moreover, the structure of the model reflected the real world situation as accurately as possible, taking into account cancelled cycles, availability of none or only one embryo for transfer, declining pregnancy rates in subsequent cycles, FET's and treatment dropouts. Also, our model reflected the general IVF population qualifying for eSET, STP or DET as accurately as possible, as both participants, eligible non-participants from a previous performed RCT (Van Montfoort *et al*., 2006b), as well as non-eligible patients were used as a basis for calculation of the most relevant probabilities in order to prevent selection bias. Seven treatment strategies were included, based on a maximum of three consecutive IVF cycles representing a similar or constructive transfer policy. Although the model is very extensive, we realize that it may not represent all possible IVF practice variations.

The Markov model showed that it was not cost-effective to combine several transfer policies (one cycle of eSET + two cycles of STP; one cycle of eSET + one cycle of STP + one cycle of DET; one cycle of eSET + two cycles of DET; one cycle of STP + two cycles of DET). Therefore, a choice has to be made between three cycles of eSET, STP or DET. It depends, however, on society's willingness to pay which strategy is to be preferred from a cost-effectiveness point of view. If this is less than €7350 for one extra live birth, three cycles of eSET are to be preferred. If society is willing to pay between €7350 and €15 250, three cycles of STP are to be preferred, and if society is willing to pay more than €15 250 per extra live birth, three cycles of DET are to be preferred. However, it should be noted that until the ceiling ratio reaches about €17 000, the probability that a strategy is to be preferred from a cost-effectiveness viewpoint is comparable between most strategies. If the ceiling ratio is higher than €17 000, three cycles of DET obviously becomes the preferred strategy. Other cost-effectiveness studies (Wolner-Hanssen and Rydhstroem, 1998; De Sutter *et al*., 2002; Gerris *et al*., 2004; Lukassen *et al*., 2005) only compared one or two cycles of eSET with one cycle of DET. One study used a modeling approach to compare eSET with DET, which was based on several assumptions, whereby it was not allowed for testing in a realistic situation (De Sutter *et al*., 2002). In another study (Wolner-Hanssen and Rydhstroem, 1998), a cohort study was combined with model calculations, and a theoretical live birth rate for eSET was used. Furthermore, an empirical study was published recently (Heijnen *et al*., 2007; Polinder *et al*., 2008), in which four cycles of eSET with mild ovarian stimulation were compared with three cycles of DET with standard stimulation. In this study, patients were younger than 38 years, had a regular menstrual cycle and were not obese. Because of these differences, results could not be compared with the results we found in our study, nor could these results be incorporated in our model. None of the studies mentioned above calculated the cost-effectiveness of several IVF strategies, like STP or a combination of several transfer policies.

The cost-effectiveness analysis presented here has some limitations that need to be addressed. First, with respect to the scope of the Markov model, the model ended if the end-point 'a live birth' was reached or if one of the other stop criteria was met. As a consequence, subsequent IVF treatments of patients with a live birth were not included in the model such as future occurrences of patients who did not become pregnant. In the IVF center of the University Hospital Maastricht, it was noted that 36% of couples who have one child will start a new IVF treatment, compared with 25% of couples who have two or more children. Therefore, it can be expected that in the eSET strategy, more patients will start a new IVF treatment for another child compared with the STP and DET strategy, as in these strategies the percentage of patients with twins is higher (so their child wish may be fulfilled). The societal consequences of these additional IVF attempts are not included in the model. Second, the input parameters for the second order Monte Carlo simulation were comparable with the parameters for the first order model, with the only difference that distributions were built in. In general, to enhance the external validity of the model, the parameter values should be determined based on a (systematic) review. However, in most studies published so far comparing eSET with DET, the population characteristics were not comparable with our unselected population of IVF patients, since they performed eSET in a selected group of patients (i.e. patients under 35–38 years of age with two or more good quality embryos) (Gerris *et al*., 2004; Lukassen *et al*., 2005; Thurin Kjellberg *et al*., 2006). One study performed both eSET and DET in a partially selected group of patients (i.e. patients under 38 years of age with a regular menstrual cycle and a BMI of 18–28 kg/m^2) (Heijnen *et al*., 2007). Therefore, it was not possible to use population parameter values emerging from a systematic review as is ideally the case in models based on evidence synthesis (Sculpher and Drummond, 2006). Parameter values were based on data from a large 'real-life' patient group originating from our IVF clinic. Nevertheless, it can be argued that the values used as input for our model may not correctly reflect the average IVF population qualifying for eSET or DET in every European country today.

The choice between three-cycle eSET, three-cycle STP or three-cycle DET strategy in our study, depends on what society is willing to pay for one extra live birth. However, it is difficult to determine which threshold value should be used because our outcome measure cannot be compared directly to commonly used outcome measures such as life years gained or quality adjusted life years (QALYs; combining the number of life years gained with the quality of that life). Nevertheless, the lesson to be learned from our study is that substituting DET by STP and STP by eSET would lead to cost savings, but also to loss of effectiveness. Whether society is willing to lose effectiveness for the individual patient remains to be seen (Severens *et al*., 2005). On the other hand, a ceiling ratio between €20 000 and €50 000 per QALY seems to be generally accepted among health-economic researchers (Hirth *et al*., 2000). As it can be argued that an average live born child will 'generate' more than one QALY, a ceiling ratio of €15 250 in this study seems low. Consequently, a threshold of more than €15 250 seems justifiable, resulting in the conclusion that the DET strategy is the strategy of choice from a cost-effectiveness point of view. This is corroborated by the study of Neumann and Johannesson (1994), who showed the willingness to pay for a child to range from €90 000 to €900 000.

By using a cost per QALY approach, it can, however, be argued that not only future effects but also future costs associated with singletons and twins should be included, to obtain a fair estimate of the incremental costs per QALY. These future costs are actually costs in life years gained, which are normally not included in cost-effectiveness analyses. Including these long-term costs in a cost-effectiveness analysis would be interesting, also from a methodological point of view (Drummond and McGuire, 2001; Van Baal *et al*., 2007). In this respect, it can be noted that recently a research project has been granted to members of our research group in which long-term costs and outcomes of IVF singletons and twins will be investigated, with the ultimate aim to incorporate these long-term consequences in our short-term Markov cost-effectiveness model.

In addition, in our base-case analysis, a twin pregnancy was defined as only one live birth. It seems to be common practice to use this outcome measure in the IVF literature, to reflect the general opinion that twins should be considered as a complication. Also from a cost per QALY point of view, twins on average will generate more QALYs compared with singletons. As 21% of the successful pregnancies in the three-cycle DET strategy were twins, including QALYs of twin births may have a considerable impact on the ICER. In this respect, it should be noticed that the additional IVF cycles, as already mentioned above, should also be included in this long-term cost per QALY analysis. It is expected that a proportion of patients, after having had a successful eSET, will make use of another IVF treatment for a second child, although they may be less likely to accept the risk of having twins. This additional treatment will generate costs for society, but children born from these additional embryo transfers will also generate QALYs. Furthermore, patients who did not become pregnant as a result of IVF often seek other options, such as adoption or foster children. The QALYs and costs of these children should also be included.

Apart from the QALYs of the children born, QALYs of the patients and family members should also be considered. For example, involuntary childlessness could have an influence on the wellbeing of the couples, and a severely handicapped child may have a considerable impact on the wellbeing of both the parents and siblings. So, for a balanced approach with respect to the long-term costs and effects of eSET, STP or DET, all these issues should be considered.

In conclusion, the Markov cost-effectiveness model presented in this paper has evaluated the full IVF procedure for a wide range of transfer policies in the general IVF population, whereby the input of the model was largely based on real-life empirical data. Our study shows that in patients qualifying for IVF treatment, combining several transfer policies was not cost-effective. A choice has to be made between three cycles of eSET, STP or DET. It depends, however, on society's willingness to pay which strategy is to be preferred from a cost-effectiveness point of view. By investigating the long-term costs and outcomes of IVF singletons and twins, we plan to give a clear answer to the question of which IVF strategy is ultimately the most cost-effective.

Supplementary data

Supplementary data are available at http://humrep.oxfordjournals.org/.

Funding

This study was supported by a research grant (number 945-12-014) from the Dutch Organisation for Health Research and Development (ZonMw) and the Dutch Health Insurance Board (CvZ) in a joint research program on health technology assessment of infertility.

References

Andersen AN, Goossens V, Gianaroli L, Felberbaum R, de Mouzon J, Nygren KG. Assisted reproductive technology in Europe, 2003. Results generated from European registers by ESHRE. *Hum Reprod* 2007;**22**:1513–1525.

Brennan A, Akehurst R. Modelling in health economic evaluation. What is its place? What is its value? *Pharmacoeconomics* 2000;**17**:445–459.

Briggs A, Claxton K, Sculpher M. *Decision-Modelling for Health Economic Evaluation*. Oxford: Oxford University Press, 2006.

Buxton MJ, Drummond MF, Van Hout BA, Prince RL, Sheldon TA, Szucs T, Vray M. Modelling in economic evaluation: an unavoidable fact of life. *Health Econ* 1997;**6**:217–227.

De Sutter P, Gerris J, Dhont M. A health-economic decision-analytic model comparing double with single embryo transfer in IVF/ICSI. *Hum Reprod* 2002;**17**:2891–2896.

Detsky AS, Naglie G, Krahn MD, Naimark D, Redelmeier DA. Primer on medical decision analysis: part 1–getting started. *Med Decis Making* 1997a;**17**:123–125.

Detsky AS, Naglie G, Krahn MD, Redelmeier DA, Naimark D. Primer on medical decision analysis: part 2–building a tree. *Med Decis Making* 1997b;**17**:126–135.

Drummond M, McGuire A. *Economic Evaluation in Health Care. Merging Theory with Practice*. Oxford, Oxford University Press, 2001.

Fenwick E, Claxton K, Sculpher M. Representing uncertainty: the role of cost-effectiveness acceptability curves. *Health Econ* 2001;**10**:779–787.

Fiddelers AA, van Montfoort AP, Dirksen CD, Dumoulin JC, Land JA, Dunselman GA, Janssen JM, Severens JL, Evers JL. Single versus double embryo transfer: cost-effectiveness analysis alongside a randomized clinical trial. *Hum Reprod* 2006;**21**:2090–2097.

Fiddelers AA, Severens JL, Dirksen CD, Dumoulin JC, Land JA, Evers JL. Economic evaluations of single- versus double-embryo transfer in IVF. *Hum Reprod Update* 2007;**13**:5–13.

Gerris J, De Sutter P, De Neubourg D, Van Royen E, Vander Elst J, Mangelschots K, Vercruyssen M, Kok P, Elseviers M, Annemans L *et al.* A real-life prospective health economic study of elective single embryo transfer versus two-embryo transfer in first IVF/ICSI cycles. *Hum Reprod* 2004;**19**:917–923.

Heijnen EM, Eijkemans MJ, De Klerk C, Polinder S, Beckers NG, Klinkert ER, Broekmans FJ, Passchier J, Te Velde ER, Macklon NS *et al*. A mild treatment strategy for in-vitro fertilisation: a randomised non-inferiority trial. *Lancet* 2007;**369**:743–749.

Hirth RA, Chernew ME, Miller E, Fendrick AM, Weissert WG. Willingness to pay for a quality-adjusted life year: in search of a standard. *Med Decis Making* 2000;**20**:332–342.

Krahn MD, Naglie G, Naimark D, Redelmeier DA, Detsky AS. Primer on medical decision analysis: part 4–analyzing the model and interpreting the results. *Med Decis Making* 1997;**17**:142–151.

Land JA, Evers JL. Risks and complications in assisted reproduction techniques: report of an ESHRE consensus meeting. *Hum Reprod* 2003;**18**:455–457.

Lukassen HG, Braat DD, Wetzels AM, Zielhuis GA, Adang EM, Scheenjes E, Kremer JA. Two cycles with single embryo transfer versus one cycle with double embryo transfer: a randomized controlled trial. *Hum Reprod* 2005;**20**:702–708.

Miller DK, Homan SM. Determining transition probabilities: confusion and suggestions. *Med Decis Making* 1994;**14**:52–58.

Naglie G, Krahn MD, Naimark D, Redelmeier DA, Detsky AS. Primer on medical decision analysis: part 3–estimating probabilities and utilities. *Med Decis Making* 1997;**17**:136–141.

Naimark D, Krahn MD, Naglie G, Redelmeier DA, Detsky AS. Primer on medical decision analysis: part 5–working with Markov processes. *Med Decis Making* 1997;**17**:152–159.

Neumann PJ, Johannesson M. The willingness to pay for in vitro fertilization: a pilot study using contingent valuation. *Med Care* 1994; **32**:686–699.

Oostenbrink J, Koopmanschap M, Rutten F. Standardisation of costs. The Dutch manual for costing in economic evaluations. *Pharmacoeconomics* 2002;**20**:443–454.

Oostenbrink J, Bouwmans C, Koopmanschap M, FFH R. Handleiding voor kostenonderzoek, methoden en standaard kostprijzen voor economische evaluaties in de gezondheidszorg. College voor zorgverzekeringen (*Guideline for Cost Research, Methods, and Standard Cost Prices for Economic Evaluations in Health Care*. Health Care Insurance Board). FO/DH&P Diemen, 2004.

Philips Z, Bojke L, Sculpher M, Claxton K, Golder S. Good practice guidelines for decision-analytic modelling in health technology assessment: a review and consolidation of quality assessment. *Pharmacoeconomics* 2006;**24**:355–371.

Polinder S, Heijnen EM, Macklon NS, Habbema JD, Fauser BJ, Eijkemans MJ. Cost-effectiveness of a mild compared with a standard strategy for IVF: a randomized comparison using cumulative term live birth as the primary endpoint. *Hum Reprod* 2008;**23**: 316–323.

Sculpher MJ, Claxton K, Drummond M, McCabe C. Whither trial-based economic evaluation for health care decision making? *Health Econ* 2006;**15**:677–687.

Sculpher MJ, Drummond MF. Analysis sans frontiers: can we ever make economic evaluations generalisable across jurisdictions? *Pharmacoeconomics* 2006;**24**:1087–1099.

Severens JL, Brunenberg DE, Fenwick EA, O'Brien B, Joore MA. Cost-effectiveness acceptability curves and a reluctance to lose. *Pharmacoeconomics* 2005;**23**:1207–1214.

Thurin Kjellberg AT, Carlsson P, Bergh C. Randomized single versus double embryo transfer: obstetric and paediatric outcome and a cost-effectiveness analysis. *Hum Reprod* 2006;**21**:210–216.

Van Baal PH, Feenstra TL, Hoogenveen RT, de Wit GA, Brouwer WB. Unrelated medical care in life years gained and the cost utility of primary prevention: in search of a 'perfect' cost-utility ratio. *Health Econ* 2007;**16**:421–433.

Van Hout BA, Al MJ, Gordon GS, Rutten FF. Costs, effects and C/E-ratios alongside a clinical trial. *Health Econ* 1994;**3**:309–319.

Van Montfoort AP, Fiddelers AA, Janssen JM, Derhaag JG, Dirksen CD, Dunselman GA, Land JA, Geraedts JP, Evers JL, Dumoulin JC. In unselected patients, elective single embryo transfer prevents all multiples, but results in significantly lower pregnancy rates compared with double embryo transfer: a randomized controlled trial. *Hum Reprod* 2006a;**21**:338–343.

Van Montfoort AP, Fiddelers AA, Janssen JM, Derhaag JG, Dirksen CD, Dunselman GA, Land JA, Geraedts JP, Evers JL, Dumoulin JC. In unselected patients, elective single embryo transfer prevents all multiples, but results in significantly lower pregnancy rates compared with double embryo transfer: a randomized controlled trial. *Hum Reprod* 2006b;**21**:338–343.

Weinstein M, O'Brien B, Hornberger J, Jackson J, Johannesson M, McCabe C, Luce B. Principles of good practice for decision analytic modeling in health-care evaluation: report of the ISPOR task force on good research practices - modeling studies. *Value health* 2003;**6**:9–17.

Wolner-Hanssen P, Rydhstroem H. Cost-effectiveness analysis of in-vitro fertilization: estimated costs per successful pregnancy after transfer of one or two embryos. *Hum Reprod* 1998;**13**:88–94.

Submitted on March 25, 2008; resubmitted on January 6, 2009; accepted on January 20, 2009