RBMO

ELSEVIER

ARTICLE

Construction and validation of a prediction model to minimize twin rates at preserved high live birth rates after IVF

BIOGRAPHY

Katarina Kebbon Vaegter has been a Specialist in Obstetrics and Gynecology since 2004, currently working at the Carl von Linné Clinic, a university-affiliated private fertility clinic in Uppsala. She is currently pursuing PhD studies into prediction models in IVF at the Department of Women's and Children's Health at Uppsala University.

Katarina Kebbon Vaegter[1,2,*], Lars Berglund[3,4], Johanna Tilly[5,1], Nermin Hadziosmanovic[3], Thomas Brodin[1,2], Jan Holte[1,2,6]

KEY MESSAGE

Part of an ongoing discussion into the potential risks of prolonged embryo culture, this study shows that with a well-constructed prediction model, an accurate choice between eSET and DET can be made as early as day 2 after oocyte retrieval, when the aim is low twin rates at preserved high live birth rates.

ABSTRACT

Research question: Elective single-embryo transfer (eSET) at blastocyst stage is widely used to reduce the frequency of multiple pregnancies after IVF. There are, however, concerns about increased risks for the offspring with prolonged embryo culture. Is it possible to select embryos for transfer at the early cleavage stage and still achieve low twin rates at preserved high live birth rates?

Design: A prediction model (PM) was developed to optimize eSET based on variables known 2 days after oocyte retrieval (fresh day 2 embryo transfers; double-embryo transfers 1999–2002 (n=2846) and SET 1999–2003 (n=945); n total=3791). Seventy-five variables were analysed for association with pregnancy chance and twin risk and combined for PM construction. This PM was validated in 2004–2016 including frozen-thawed transfers (FET), to compare cumulative live birth rate (CLBR) and twin rate before (1999–2002 fresh embryo transfers plus FET from the same oocyte retrievals until the end of 2007, n=3495) and after (2004–2011 fresh embryo transfers plus FET from the same oocyte retrievals until the end of 2016, n=11195) implementing the model.

Results: The PM was constructed from four independent variables: female age, embryo score, ovarian sensitivity and treatment history. The calibration, i.e. the fit of observed versus predicted results, was excellent both at construction and at validation. Without compromising CLBR, twin rate was reduced from 25.2% to 3.8%, accompanied by profound improvements in perinatal outcome.

Conclusion: The results provide the first successful construction, validation and impact analysis of a day 2 transfer PM to reduce multiple pregnancies.

[1] Department of Women's and Children's Health, Uppsala University, Akademiska Sjukhuset, Uppsala S-751 85, Sweden
[2] Carl von Linné Clinic, Uppsala Science Park, Uppsala S-751 83, Sweden
[3] Uppsala Clinical Research Centre, Uppsala University, Uppsala S-751 85, Sweden
[4] Department of Public Health and Caring Sciences/Geriatrics, Uppsala University, Sweden
[5] Department of Mathematics, Uppsala University, Uppsala S-751 06, Sweden
[6] Centre for Reproductive Biology in Uppsala, University of Agricultural Science and Uppsala University, Uppsala S-750 07, Sweden
[1]Present address: PCG Clinical Services, S-753 23 Uppsala, Sweden

KEYWORDS

Double-embryo transfer (DET)
IVF/ICSI outcome
Multiple pregnancy
Prediction model
Single-embryo transfer (SET)

*Corresponding author. K. Kebbon Vaegter. E-mail addresses: katarina.kebbon_vaegter@kbh.uu.se, katarina.vaegter@linne.se (K K Vaegter). https://doi.org/10.1016/j.rbmo.2018.09.020 1472-6483/
Declaration: The authors report no financial or commercial conflicts of interest.

INTRODUCTION

Twin pregnancies suffer from greatly increased risks, independent of mode of conception. The incidence of premature birth, low birth weight, cerebral palsy and other neurological complications, as well as perinatal mortality, is markedly increased compared with singleton pregnancies (*Bergh et al., 1999*). In the USA in 2012, the multiple birth rate after IVF was 27%. The overall twin rate in Europe after assisted reproduction is decreasing but is still 17% (*Calhaz-Jorge et al., 2016*). As double-embryo transfer (DET) strategies result in twin pregnancies at frequencies of 25–30% in normal IVF populations, and 1.6% of babies born in the USA (2012) and 1–6% of all children in European countries nowadays result from IVF, the hazards of these iatrogenic twin pregnancies cause considerable suffering for many couples and also huge extra costs for society (*Calhaz-Jorge et al., 2016; Luke et al., 2015*). In order to reduce multiple births after assisted reproductive technology (ART), elective single-embryo transfer (eSET) strategies need to be optimized. Single-embryo transfer (SET) at the blastocyst stage has become the strategy of choice for many clinics worldwide, because blastocyst transfer improves the odds of transferring a viable embryo (*Harton et al., 2013*). However, reports on a higher incidence of preterm delivery, monozygotic twins, large for gestational age babies, congenital anomalies and altered sex ratio with blastocyst transfers compared with cleavage-stage transfers raise concerns. Moreover, possible epigenetic changes resulting from the prolonged culture have been discussed (*Chang et al., 2009; Kallen et al., 2010b; Maheshwari et al., 2016*). Extending the culture to blastocyst stage is also associated with an increased risk of cycle cancellation due to lack of embryos to transfer. An accurate choice between eSET and DET at the cleavage stage would therefore be preferable.

Clearly, eSET in exclusively good-prognosis treatments generally yield acceptable pregnancy rates, especially if the increased number of frozen-thawed embryo transfers (FET) is considered (*Lundin and Bergh, 2007; Saldeen and Sundstrom, 2005; Strandell et al., 2000; Thurin et al., 2004*). However, in normal IVF populations a large proportion of patients and treatments must be regarded as sub-optimal in terms of prognosis. Thus, the difficulty is choosing between SET and DET when the embryo morphology is sub-optimal, when the couple has failed in previous attempts, and/or when the woman is of greater age or responds poorly to ovarian stimulation. In order to optimize the balance between a low rate of multiple pregnancies and an overall acceptable and high pregnancy rate, an increased knowledge of the factors determining implantation after ART is needed. The purpose of this series of studies was to establish and validate algorithms for implantation potential and twin risk in patients covering the full range of prognostic potential in an ordinary IVF population. Subsequently, these algorithms formed a prediction model (PM) that was applied in all treatments, aimed at radically reducing the twin rate. Birth rates, twin rates and perinatal data were compared between periods before and after applying the PM. This constitutes the final step in the evolution of a PM – an impact analysis – that evaluates whether the model improves decisions in terms of quality or effectiveness of patient care (*Leushuis et al., 2009; van Loendersloot et al., 2014*).

MATERIALS AND METHODS

Data from all IVF/intracytoplasmic sperm injection (ICSI) treatments were recorded prospectively from 1999. The couples all had an infertility duration of at least 1 year, and had gone through an infertility investigation. The vast majority were Swedish Caucasians. Treatments could be either private or government-funded. Treatments up to 2016 resulting in embryo transfer on day 2 after oocyte retrieval were included as follows.

Construction of a DET and twin algorithm

All fresh IVF/ICSI treatments that resulted in DET during 1999–2002 ('2ET') (n = 2846) were used to construct the DET algorithm for the chance of a clinical pregnancy after DET and the twin algorithm for the chance of a twin pregnancy after DET.

Construction of a SET algorithm

All fresh IVF/ICSI treatments that resulted in SET from 1999 to 2003 ('1ET') (n = 945) were used together with '2ET' to construct the SET algorithm for the chance of a clinical pregnancy after SET. (I) and (II) together form the PM for guidance of eSET versus DET.

Validation of the model

From 2004 onwards, the PM has guided the selection of eSET or DET in individual cases, and in the vast majority of cases (88%), eSET and DET were performed accordingly. All IVF/ICSI treatments during 2004–2011 resulting in fresh embryo transfer on day 2 after oocyte retrieval constitute the validation dataset (n = 7515). The ability of the model to discriminate couples that would become pregnant and/or have a twin pregnancy from those that would not ['discrimination', expressed as the area under the receiver operating characteristic curve (AUC), or c-statistics] and the concordance between the observed and the predicted results for pregnancy rates and twin rates ('calibration') were estimated.

Analysis of transfers not following the PM

The minority of treatments (n = 879) with a number of embryos to embryo transfer not following the suggestion from the PM were analysed separately. Specifically, the smaller number of DET (n = 278) performed in spite of an estimated twin risk above 15% enabled a validation of the twin algorithm, albeit the number of treatments was small and hence the statistical power limited (see below).

Impact analysis

All fresh day 2 embryo transfers in 1999–2002, plus FET cycles from the same oocyte retrieval up to the end of 2007 ('Before PM', n = 3495), were compared with all fresh embryo transfers during the validation period (2004–2011), plus FET cycles from the same oocyte retrieval up to the end of 2016 ('With PM', n = 11,195), for live birth rate (LBR), twin rate, cumulative live birth rate (CLBR) and perinatal outcome. CLBR was defined as the first live birth following either a fresh embryo transfer or a FET from the same oocyte retrieval.

Ovarian stimulation

Ovarian stimulation was conducted as previously described (*Brodin et al., 2009, 2013*). During the 'Before PM' period 98% of the cycles were agonist protocols and 2% were antagonist protocols, with recombinant FSH (rFSH) in 99% of the cycles and highly purified human menopausal gonadotrophin (HMG) in 1%. During the validation period 89% of the cycles were agonist protocols and 11% were antagonist protocols, with rFSH in 72% of the cycles and HMG in

28%. Oocytes were aspirated 36–37 h after administration of hCG. Fertilization, embryo evaluation and embryo transfer were performed as described previously (*Holte et al., 2007*).

Embryo transfer strategies

During 1999–2002, the general policy was to transfer two embryos when available, mainly independent of prognosis and twin risk, i.e. following the most common policy in Europe at the time. During this period, FET followed the same strategy. Thus SET amounted to only 11.1% and the majority of those SET were not elective, i.e. only one embryo was available for transfer. From 2004 the following eSET policy was applied: if the PM suggested a risk of twin implantation above 15% using the two highest scored embryos, then only one embryo was transferred. If the model suggested a lower twin risk than 15%, the couple was offered two embryos for transfer (if the medical history did not constitute any contraindications). According to data from the construction period, about 30% of the couples would be offered DET, resulting in a predicted twin rate of 0–15%, i.e. the twin rate in that group would be approximately 7.5%. The predicted twin implantation rate in the entire population would thus amount to 7.5% × 0.30, i.e. around 2–3%. All surplus embryos of high morphological score were cryopreserved by slow freezing; there was no use of vitrification. Selection of frozen-thawed embryos for eSET or DET generally followed the same criteria as for the corresponding fresh treatment, i.e. eSET was performed during the validation period if this had been performed in the fresh treatment and the embryo score of the thawed embryo did not change this decision.

Statistical methods

The aim of the statistical analyses was to estimate algorithms for clinical pregnancy and for twins given pregnancy when two embryos were transferred, and for clinical pregnancy when one embryo was transferred. Estimates were based on selected predictors. The individual twin probability was to be used as a tool to decide whether one or two embryos should be offered for transfer.

Estimation of pregnancy chance and twin risk for DET

To estimate probabilities for pregnancy and for twins given pregnancy when two embryos were transferred, data from 2846 DET treatments from 1999 to 2002 were used. The data set was randomly divided into two groups, one training data set (TDS), which constituted two-thirds of the observations, and a validation data set (VDS, the remaining one-third of the observations). The latter was used to examine the predictive capacity of the final algorithms. Seventy-five putative predictors such as the woman's age, number of treatments, number of previous pregnancies, infertility cause, weight, FSH dose and embryo scores (Supplemental 1) were used. The outcome was the number of clinical pregnancies per embryo transfer (ultrasonographically verified presence of 0, 1 or 2 gestational sacs). The outcome was dichotomized as 1 or 2 versus 0 number of sacs in the pregnancy chance algorithm and as 2 versus 1 (excluding 0) number of sacs in the algorithm for twin risk given pregnancy.

Use of TDS. Logistic regression analyses were used to identify and summarize combinations of predictors to be used in the final algorithms. Predictors with a univariate *P*-value <0.1 were selected as predictors in multivariable models. The use of composite variables, categorization of continuous variables (e.g. woman's age), as well as interaction terms of the second degree, were allowed in the multivariable models. Significance criteria for predictors in the multivariable models were $P < 0.05$.

Use of VDS. The estimated odds ratios (OR) and c-statistics from TDS were compared with the corresponding statistics from VDS and if all estimated statistics had an absolute difference of less than 10% of the value from TDS the two data sets were merged and all statistics were estimated from the complete data set.

The complete data set. The final multivariable algorithms for pregnancy and for twins given pregnancy, were presented with OR with 95% confidence intervals (CI), *P*-values and c-statistics.

Estimation of pregnancy chance for SET (Supplemental 2)

Variables. All putative predictors are shown in Supplemental. Supplemental 3 shows the variables that univariately correlated with clinical pregnancy rate (CPR). The integrated morphology cleavage (IMC) embryo score is an evidence-based embryo scoring model for embryos on day 2 after oocyte retrieval (*Holte et al., 2007*). 'Treatment history' is a composite variable composed of the number of earlier treatments and any resulting pregnancies. 'Ovarian sensitivity' is also a composite variable describing the number of retrieved oocytes in relation to the administered total dose of FSH (*Huber et al., 2013*).

Validation statistics

The discriminative capacity of the algorithms was described as the c-statistics (or AUC). Calibration of the algorithms was analysed by comparing the predicted CPR in 10% strata with the observed CPR. The corresponding analysis was performed for twin rates in 15% strata (due to the low number of twins during the validation period). Calibration was assessed by means of the Hosmer–Lemeshow test. A non-significant *P*-value ($P > 0.05$) in the Hosmer–Lemeshow test indicates a good concordance between the predicted and the observed chance/risk.

Ethical approval. The Regional Ethics Committee at Uppsala University approved the study (2012-07-05; Dnr 2012/036) and waived the need for written informed consents.

RESULTS

Construction of the DET/SET model (1999–2003)

Twenty-nine variables out of the recorded 75 were univariately correlated with CPR in the DET material (Supplemental 3). After logistic regression analyses, the following variables remained significant: woman's age, IMC embryo score (for DET: the highest score and the difference between the scores of the two transferred embryos), the treatment history, and the ovarian sensitivity. TABLE 1 shows the OR for these variables.

Variables associated with twin implantation were analogous to those correlated to CPR (data not shown) and a similar algorithm was derived after logistic regression analyses (TABLE 1). The observed twin rate was higher than might be expected if the implantation chances of the embryos transferred together were completely independent of one another. Also, this interdependence varied in different prognostic strata (Supplemental 2). The algorithm for

TABLE 1 VARIABLES INCLUDED IN THE ALGORITHMS FOR PREGNANCY CHANCE WITH SET AND DET AND TWIN RISK, WHICH WERE USED IN THE PREDICTION MODEL

	DET pregnancy chance		Twin risk		SET pregnancy chance	
Predictor	**OR (95% CI)**	***P*-value**	**OR (95% CI)**	***P*-value**	**OR (95% CI)**	***P*-value**
Age group[a]	0.87 (0.82–0.91)	<0.0001	0.78 (0.70–0.88)	<0.0001	0.87 (0.82–0.92)	<0.0001
Treatment history[b]	0.71 (0.61–0.84)	<0.0001	0.72 (0.54–0.97)	0.03	0.71 (0.60–0.83)	<0.0001
Ovarian sensitivity[c]	0.71 (0.63–0.80)	<0.0001	0.75 (0.59–0.94)	0.01	0.70 (0.62–0.78)	<0.0001
IMC embryo score[d] highest	2.12 (1.82–0.46)	<0.0001			2.16 (1.86–2.51)	<0.0001
IMC embryo score difference	0.77 (0.68–0.88)	<0.0001				
IMC embryo score sum			1.52 (1.31–1.76)	<0.0001		

DET = double embryo transfer; IMC = integrated morphology cleavage (embryo score); SET = single embryo transfer.

[a] Age was grouped in seven groups from ≤29 to ≥42 years

[b] Treatment history is a composite variable based on any earlier IVF treatment results; two groups were formed.

[c] Ovarian sensitivity (three groups) is a composite variable (total number of eggs/total dose of FSH administered).

[d] IMC embryo score (Holte *et al.*, 2007) is an evidence-based embryo score, here sub-grouped into five groups.

pregnancy chance after eSET included the same variables as the DET algorithm, except that only one embryo was included (TABLE 1). The algorithms derived from the TDS were subsequently tested on the VDS, and because they yielded largely similar results, the two data sets were merged and all statistics were estimated from the complete data set.

Validation of the model (2004–2011)

The discriminating capacity of the algorithms was moderate, with c-statistics between 0.64 and 0.75 (Supplemental 4). The concordance between the predicted and the observed results was excellent, as shown in FIGURE 1. The CPR for each 10% stratum calibrated well, confirmed by Hosmer–Lemeshow test ($P = 0.08$). Twin rates were the expected, given the acceptance level of 15% in each case.

Analysis of transfers not following the PM

In 278 cases DET was performed even though the twin risk was >15% (physician's decision), and this resulted in the predicted twin rate >15%, although the Hosmer–Lemeshow test could not confirm a perfect calibration ($P = 0.03$). In 601 cases SET was performed despite a twin risk below 15%, for medical reasons or at the request of the couple, resulting in the predicted CPR for SET, which was lower than the predicted CPR for DET (data not shown).

Impact analysis

The populations during the 'Before PM' period and the 'With PM' period had slightly different demographic profiles (TABLE 2). During the 'With PM' period mean age was higher, the couples had performed more treatments on average, and the mean embryo score was slightly higher. LBR, CLBR and perinatal outcome data from the two periods, i.e. before and after applying the PM, are presented in TABLE 3. The twin rate was reduced from 25.2% to 3.8% after applying the model. LBR per fresh embryo transfer was lower after applying the model, but the CLBR (i.e. including FET) was marginally higher. If groups were age-adjusted this difference became highly significant ($P < 0.0001$). When adjusting the populations according to all the variables in the PM, i.e. age, embryo score, treatment history and

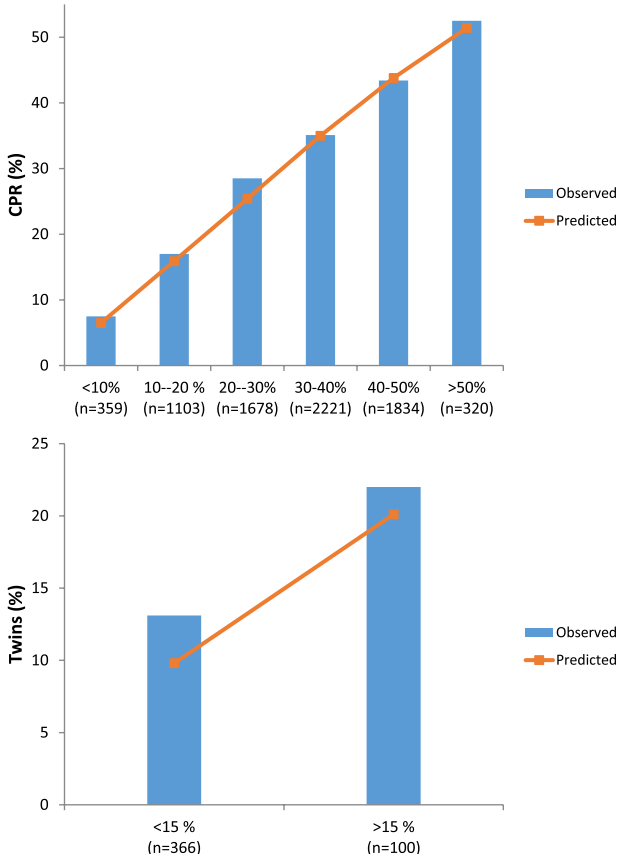

FIGURE 1 Calibration of the prediction model (fresh treatments). Observed CPR versus predicted CPR (upper panel), and observed twin rate versus predicted twin rate (lower panel), during the validation period (CPR = clinical pregnancy rate).

TABLE 2 DESCRIPTION OF THE TWO STUDY POPULATIONS IN THE IMPACT ANALYSIS, BASED ON THE VARIABLES THAT QUALIFIED IN THE PREDICTION MODEL, AND BMI

Variable	Before PM (1999–2002)	With PM (2004–2011)	*P*-value
Number of fresh ET	3163	7515	
Woman's age	34.1 (33.9–34.2)	35.0 (34.9–35.1)	<0.0001
BMI	23.5 (23.4–23.6)	23.4 (23.3–23.5)	NS
Mode of fertilization			NS
IVF	1744 (55)	4074 (54)	
ICSI	1077 (34)	2706 (36)	
Combined	342 (11)	730 (10)	
Main infertility diagnosis			<0.0001
Tubal factor	657 (21)	908 (13)	
Unexplained	1108 (35)	3487 (48)	
Male	962 (31)	1908 (26)	
Endometriosis	192 (6)	371 (5)	
Anovulation	208 (7)	519 (7)	
Other	21 (0.7)	34 (0.5)	
Eggs at oocyte retrieval	11.0 (10.8–11.2)	9.5 (9.4–9.6)	<0.0001
Total dose of FSH (IU)	2473 (2430–2516)	2451 (2321–2481)	NS
Ovarian sensitivity[a]	1.74 (1.71–1.77)	1.77 (1.75–1.78)	NS
Previous IVF treatment	0.31 (0.28–0.34)	1.35 (1.31–1.39)	<0.0001
Previous IVF pregnancy	0.09 (0.08–0.10)	0.36 (0.34–0.37)	<0.0001
Treatment history[b]	1.37 (1.35–1.39)	1.33 (1.32–1.34)	<0.0001
IMC embryo score, 1–10	8.80 (8.74–8.86)	9.11 (9.07–9.14)	<0.0001

Values are mean (95% CI) or *n* (%), unless otherwise stated.

BMI = body mass index; CI = confidence interval; ET = embryo transfer; FSH = follicle-stimulating hormone; IU = international units; IMC = integrated morphology cleavage; OPU = oocyte pick-up; PM = prediction model.

[a] The composite variable 'Ovarian sensitivity' with its groups was formed by cross-tabulation of total number of eggs and total dose of FSH.

[b] The composite variable 'Treatment history' with its groups was formed by cross-tabulation of earlier IVF treatments and children after IVF.

ovarian sensitivity, both LBR and CLBR were higher during the 'With PM' period ($P = 0.012$ and $P < 0.001$, respectively). Mean birth weights increased from 3107 ± 793 g to 3403 ± 637 g ($P < 0.0001$) after PM introduction. With the PM, the frequencies of babies born prematurely (before gestational week 33) and babies born with a low birth weight (below 2500 g) were reduced by two-thirds ($P < 0.0001$). Also, after implementing the PM, the frequency of babies born small for gestational age (SGA) was reduced by 60% ($P < 0.001$) and perinatal mortality was also lower after adjustment for age (TABLE 3).

DISCUSSION

This is thought to be the first study showing the construction, validation and subsequent impact analysis of a PM for clinical pregnancy chance after SET and DET, and twin risk following DET in IVF/ICSI. In addition, applying the model with the radical aim of reducing twin rates to <5% was successful, without compromising overall delivery rates and with dramatically improved neonatal outcome.

Largely in line with previous studies on PMs in ART, the determining variables included the age of the woman, the embryo score, ovarian sensitivity and information on treatment history (*Cai et al., 2011; Choi et al., 2013; Hunault et al., 2002; Luke et al., 2014; McLernon et al., 2016; Nelson and Lawlor, 2011; Ottosen et al., 2007; Templeton et al., 1996; van Loendersloot et al., 2013*). These variables (with slight variations) were also part of a recently published PM for live birth after SET (*Vaegter et al., 2017*).

Importantly, these results were not compatible with a simple binomial distribution of twins, singletons and failed implantation after DET. On the contrary, the embryos exhibited statistical interdependence, varying depending on the overall prognostic level. This should not be interpreted as a direct interactive effect between the embryos, which is unlikely. A more plausible interpretation is that this phenomenon reflects the sum of important patient and cycle covariates not measured (or even possible to measure), which are common to the embryos in terms of implantation conditions for the specific cycle. This principle was previously discussed in an embryo–uterus modelling framework (*Hunault et al., 2002; Roberts et al., 2010a, 2010b, 2011*).

The calibration of the model showed a high concordance between predicted and observed CPR for each 10% stratum. The modest discrimination is in accordance with previous PM in ART (*Choi et al., 2013; Leushuis et al., 2009; van Loendersloot et al., 2013*), and also represents a weakness in the present model, as it affects the accuracy of the decision making in the individual case. However, calibration over the entire range of CPR during the validation period

TABLE 3 IMPACT ANALYSIS. OUTCOME BEFORE ('BEFORE PM') AND AFTER ('WITH PM') APPLYING THE PREDICTION MODEL

Variable	'Before PM'	'With PM'	OR (95% CI)	*P*-value	Age-adjusted *P*-value	Model-adjusted *P*-value
SET (%)	11.3	75.5		<0.0001		
DET (%)	88.7	24.5		<0.0001		
LBR[a]	29.0 (27.4–30.6)	25.1 (24.1–26.0)		<0.0001	NS	0.012
CLBR[b]	30.6 (29.0–32.2)	32.6 (31.6–33.7)		NS	<0.0001	<0.001
FET[c]	9.0 (8.0–10.0)	32.8 (31.7–33.8)		<0.0001		
Twins	25.2 (22.5–28.0)	3.8 (3.0–4.5)		<0.0001		
Caesarean	32.2 (29.3–35.1)	28.0 (26.3–29.7)		0.015	0.0010	
Birth weight (g)	3107 (3062–2151)	3403 (3379–3427)		<0.0001	<0.0001	
<2500	20.3 (18.0–22.6)	7.7 (6.7–8.7)	3.06 (2.34–3.85)	<0.0001	<0.0001	
<1500	3.7 (2.6–4.7)	1.2 (0.8–1.6)	3.21 (1.92–5.35)	<0.0001	<0.0001	
<37 weeks	17.1 (14.7–19.5)	8.3 (7.2–9.4)	2.27 (1.83–2.83)	<0.0001	<0.0001	
<33 weeks	4.8 (3.5–6.2)	1.6 (1.1–2.0)	3.22 (2.10–4.92)	<0.0001	<0.0001	
SGA	2.8 (1.9–3.8)	1.1 (0.7–1.5)	2.59 (1.55–4.31)	<0.001	<0.001	
Perinatal mortality	1.0 (0.5–1.6)	0.5 (0.2–0.8)	2.06 (0.94–4.51)	NS	0.045	

Values are % (95% CI) unless otherwise stated.

CI = confidence interval; CLBR = cumulative live birth rate; DET = double embryo transfer; FET = frozen-thawed embryo transfer; LBR = live birth rate; OR = odds ratio; PM = prediction model; SET = single embryo transfer; SGA = small for gestational age.

[a] LBR is live birth rate after a fresh cycle.

[b] CLBR is the first live birth after either a fresh or a frozen-thawed cycle.

[c] FET (%) is the rate of fresh cycles that was followed by one or several frozen-thawed cycles.

shows that the model is robust and highly effective at a group level.

The frequency of SGA was more than halved. Applying the model did not compromise the overall LBR, as the slight reduction in LBR in fresh cycles was compensated for when adding frozen-thawed cycles from the same oocyte retrieval. This is well in line with the paper from *Luke et al. (2015)* that, based on a large data set from the USA, showed that the cumulative LBR over two cycles (two fresh cycles or one fresh and one frozen-thawed cycle) with SET was similar to or better than the LBR with DET in a single cycle, while the probability of multiple birth was reduced by over 90%.

The grounds for transferring embryos on day 2 rather than day 3 was to keep the culture period as short as possible. Because concerns have been raised about prolonged embryo culture, with a possible risk of epigenetic modification and potential increased risks to fetal health (*Chang et al., 2009; Dar et al., 2014; Kallen et al., 2010b; Luke et al., 2014; Luna et al., 2007; Maheshwari et al., 2016; Zhu et al., 2014*), it is an important finding that with a well-constructed PM an accurate choice between eSET and DET can be made as early as day 2. Currently there is no high-quality evidence to support the use of either blastocyst transfer or cleavage-stage transfer when the CLBR is considered (*Glujovsky et al., 2016*). Extended culture is also associated with an increased risk of having no embryos to transfer or freeze (*Quea et al., 2007*). There are still some patients for whom a cleavage-stage transfer remains the better option for optimizing the LBR (*Goldman et al., 2016*). However, (single) blastocyst culture is the current trend and to get a useful and updated PM, we are now working on an embryo scoring system for blastocysts to be incorporated in a PM for transfer on day 5.

Shortcomings of the present study include comparing results from two different time periods. IVF success rates are generally improving slightly over time, as is perinatal care. However, a randomized design was not possible for legislative reasons after a 'low twin frequency legislation' in Swedish ART was introduced in 2003 (SOSFS 2002:13, 9 kap; 1§). The overall LBR has remained stable after this policy change according to national reports (*Kallen et al., 2010,2010a; Saldeen and Sundstrom, 2005*). In a national survey, in comparison with clinics applying arbitrary grounds for eSET or DET, the use of the present PM resulted in a lower multiple birth rate without compromising the overall LBR (*Karlstrom and Bergh, 2007*).

In summary, we present the construction, validation and impact analysis of a PM for eSET or DET in the early cleavage stage, covering the entire range of a normal ART population in terms of age and treatment prognosis. It was found that the woman's age, the embryo score, ovarian sensitivity and rank and history of previous treatments together predict implantation and twin risk, and these factors formed the prediction algorithms. Applying the model with the aim of reducing twin pregnancies to below 5% resulted in the expected decrease in twin rates, without compromising the overall LBR per oocyte retrieval, followed by a marked improvement in perinatal outcome. The calibration of the model was good both for predicting pregnancy and twin implantation. We suggest that PMs like this may solve the problem of accurately selecting for eSET or DET to optimize the delicate balance between high LBR and low twin rates, the net result being a preserved high overall success rate and an improved perinatal outcome.

ACKNOWLEDGEMENTS

The authors thank the participating couples as well as all staff at the Carl von Linné Clinic. Thanks to Helena Pettersson, Department of Mathematics, Uppsala University, presently at UCR, for constructive work with the SET algorithm.

SUPPLEMENTARY MATERIALS

Supplementary material associated with this article can be found, in the online version, at doi: 10.1016/j.rbmo.2018.09.020.

REFERENCES

Bergh, T., Ericson, A., Hillensjo, T., Nygren, K.G., Wennerholm, U.B. **Deliveries and children born after in-vitro fertilization in Sweden 1982–95: a retrospective cohort study.** Lancet 1999; 354: 1579–1585

Brodin, T., Bergh, T., Berglund, L., Hadziosmanovic, N., Holte, J. **High basal LH levels in combination with low basal FSH levels are associated with high success rates at assisted reproduction.** Hum Reprod 2009; 24: 2755–2759

Brodin, T., Hadziosmanovic, N., Berglund, L., Olovsson, M., Holte, J. **Antimullerian hormone levels are strongly associated with live-birth rates after assisted reproduction.** J Clin Endocrinol Metab 2013; 98: 1107–1114

Cai, Q.F., Wan, F., Huang, R., Zhang, H.W. **Factors predicting the cumulative outcome of IVF/ICSI treatment: a multivariable analysis of 2450 patients.** Hum Reprod 2011; 26: 2532–2540

Calhaz-Jorge, C., de Geyter, C., Kupka, M.S., de Mouzon, J., Erb, K., Mocanu, E., Motrenko, T., Scaravelli, G., Wyns, C., Goossens, V. **the European IVF-monitoring (EIM) Consortium for the European Society of Human Reproduction and Embryology (ESHRE). Assisted reproductive technology in Europe, 2012: results generated from European registers by ESHRE.** Hum Reprod 2016; 31: 1638–1652

Chang, H.J., Lee, J.R., Jee, B.C., Suh, C.S., Kim, S.H. **Impact of blastocyst transfer on offspring sex ratio and the monozygotic twinning rate: a systematic review and meta-analysis.** Fertil Steril 2009; 91: 2381–2390

Choi, B., Bosch, E., Lannon, B.M., Leveille, M.C., Wong, W.H., Leader, A., Pellicer, A., Penzias, A.S., Yao, M.W. **Personalized prediction of first-cycle *in vitro* fertilization success.** Fertil Steril 2013; 99: 1905–1911

Dar, S., Lazer, T., Shah, P.S., Librach, C.L. **Neonatal outcomes among singleton births after blastocyst versus cleavage stage embryo transfer: a systematic review and meta-analysis.** Hum Reprod Update 2014; 20: 439–448

Glujovsky, D., Farquhar, C., Quinteiro Retamar, A.M., Alvarez Sedo, C.R., Blake, D. **Cleavage stage versus blastocyst stage embryo transfer in assisted reproductive technology.** Cochrane Database Syst Rev 2016 CD002118

Goldman, R.H., Kaser, D.J., Missmer, S.A., Srouji, S.S., Farland, L.V., Racowsky, C. **Building a model to increase live birth rate through patient-specific optimization of embryo transfer day.** J Assist Reprod Genet 2016; 33: 1525–1532

Harton, G.L., Munne, S., Surrey, M., Grifo, J., Kaplan, B., McCulloh, D.H., Griffin, D.K., Wells, D., Practitioners Group, P.G.D. **Diminished effect of maternal age on implantation after preimplantation genetic diagnosis with array comparative genomic hybridization.** Fertil Steril 2013; 100: 1695–1703

Holte, J., Berglund, L., Milton, K., Garello, C., Gennarelli, G., Revelli, A., Bergh, T. **Construction of an evidence-based integrated morphology cleavage embryo score for implantation potential of embryos scored and transferred on day 2 after oocyte retrieval.** Hum Reprod 2007; 22: 548–557

Huber, M., Hadziosmanovic, N., Berglund, L., Holte, J. **Using the ovarian sensitivity index to define poor, normal, and high response after controlled ovarian hyperstimulation in the long gonadotrophin-releasing hormone-agonist protocol: suggestions for a new principle to solve an old problem.** Fertil Steril 2013; 100: 1270–1276

Hunault, C.C., Eijkemans, M.J., Pieters, M.H., te Velde, E.R., Habbema, J.D., Fauser, B.C., Macklon, N.S. **A prediction model for selecting patients undergoing *in vitro* fertilization for elective single embryo transfer.** Fertil Steril 2002; 77: 725–732

Kallen, A.J., Finnstrom, O.O., Lindam, A.P., Nilsson, E.M., Nygren, K.G., Olausson, P.M. **Cerebral palsy in children born after *in vitro* fertilization. Is the risk decreasing?.** Eur J Paediatr Neurol 2010; 14: 526–530

Kallen, B., Finnstrom, O., Lindam, A., Nilsson, E., Nygren, K.G., Olausson, P.O. **Blastocyst versus cleavage stage transfer in *in vitro* fertilization: differences in neonatal outcome?.** Fertil Steril 2010; 94: 1680–1683

Kallen, B., Finnstrom, O., Lindam, A., Nilsson, E., Nygren, K.G., Otterblad Olausson, P. **Trends in delivery and neonatal outcome after *in vitro* fertilization in Sweden: data for 25 years.** Hum Reprod 2010; 25: 1026–1034

Karlstrom, P.O., Bergh, C. **Reducing the number of embryos transferred in Sweden-impact on delivery and multiple birth rates.** Hum Reprod 2007; 22: 2202–2207

Leushuis, E., van der Steeg, J.W., Steures, P., Bossuyt, P.M., Eijkemans, M.J., van der Veen, F., Mol, B.W., Hompes, P.G. **Prediction models in reproductive medicine: a critical appraisal.** Hum Reprod Update 2009; 15: 537–552

Luke, B., Brown, M.B., Wantman, E., Stern, J.E. **Factors associated with monozygosity in assisted reproductive technology pregnancies and the risk of recurrence using linked cycles.** Fertil Steril 2014; 101: 683–689

Luke, B., Brown, M.B., Wantman, E., Stern, J.E., Baker, V.L., Widra, E., Coddington, C.C., 3rd, W.E.Gibbons, Ball, G.D. **A prediction model for live birth and multiple births within the first three cycles of assisted reproductive technology.** Fertil Steril 2014; 102: 744–752

Luke, B., Brown, M.B., Wantman, E., Stern, J.E., Baker, V.L., Widra, E., Coddington, C.C.3rd, Gibbons, W.E., Van Voorhis, B.J., Ball, G.D. **Application of a validated prediction model for *in vitro* fertilization: comparison of live birth rates and multiple birth rates with 1 embryo transferred over 2 cycles versus 2 embryos in 1 cycle.** Am J Obstet Gynecol 2015; 212: e1–e7

Luna, M., Duke, M., Copperman, A., Grunfeld, L., Sandler, B., Barritt, J. **Blastocyst embryo transfer is associated with a sex-ratio imbalance in favour of male offspring.** Fertil Steril 2007; 87: 519–523

Lundin, K., Bergh, C. **Cumulative impact of adding frozen-thawed cycles to single versus double fresh embryo transfers.** Reprod Biomed Online 2007; 15: 76–82

Maheshwari, A., Hamilton, M., Bhattacharya, S. **Should we be promoting embryo transfer at blastocyst stage?.** Reprod Biomed Online 2016; 32: 142–146

McLernon, D.J., Steyerberg, E.W., Te Velde, E.R., Lee, A.J., Bhattacharya, S. **Predicting the chances of a live birth after one or more complete cycles of *in vitro***

fertilization: population based study of linked cycle data from 113 873 women. BMJ 2016; 355: i5735

Nelson, S.M., Lawlor, D.A. **Predicting live birth, preterm delivery, and low birth weight in infants born from *in vitro* fertilization: a prospective study of 144,018 treatment cycles.** PLoS Med 2011; 8: e1000386

Ottosen, L.D., Kesmodel, U., Hindkjaer, J., Ingerslev, H.J. **Pregnancy prediction models and eSET criteria for IVF patients–do we need more information?.** J Assist Reprod Genet 2007; 24: 29–36

Quea, G., Romero, K., Garcia-Velasco, J.A. **Extended embryo culture to increase implantation rate.** Reprod Biomed Online 2007; 14: 57–65

Roberts, S.A., Hirst, W.M., Brison, D.R., Vail, A. **S. E. T. collaboration toward. Embryo and uterine influences on IVF outcomes: an analysis of a UK multi-centre cohort.** Hum Reprod 2010; 25: 2792–2802

Roberts, S.A., McGowan, L., Mark Hirst, W., Vail, A., Rutherford, A., Lieberman, B.A., Brison, D.R. **S. E. T. Collaboration toward. Reducing the incidence of twins from IVF treatments: predictive modelling from a retrospective cohort.** Hum Reprod 2011; 26: 569–575

Roberts, S., McGowan, L., Hirst, W., Brison, D., Vail, A., Lieberman, B. **Towards single embryo transfer? Modelling clinical outcomes of potential treatment choices using multiple data sources: predictive models and patient perspectives.** Health Technol Assess 2010; 14: 1–237

Saldeen, P., Sundstrom, P. **Would legislation imposing single embryo transfer be a feasible way to reduce the rate of multiple pregnancies after IVF treatment?.** Hum Reprod 2005; 20: 4–8

Strandell, A., Bergh, C., Lundin, K. **Selection of patients suitable for one-embryo transfer may reduce the rate of multiple births by half without impairment of overall birth rates.** Hum Reprod 2000; 15: 2520–2525

Templeton, A., Morris, J.K., Parslow, W. **Factors that affect outcome of in-vitro fertilization treatment.** Lancet 1996; 348: 1402–1406

Thurin, A., Hausken, J., Hillensjo, T., Jablonowska, B., Pinborg, A., Strandell, A., Bergh, C. **Elective single-embryo transfer versus double-embryo transfer in *in vitro* fertilization.** N Engl J Med 2004; 351: 2392–2402

Vaegter, K.K., Lakic, T.G., Olovsson, M., Berglund, L., Brodin, T., Holte, J. **Which factors are most predictive for live birth after *in vitro* fertilization and intracytoplasmic sperm injection (IVF/ICSI) treatments? Analysis of 100 prospectively recorded variables in 8,400 IVF/ICSI single-embryo transfers.** Fertil Steril 2017; 107: 641–648

van Loendersloot, L.L., van Wely, M., Repping, S., Bossuyt, P.M., van der Veen, F. **Individualized decision-making in IVF: calculating the chances of pregnancy.** Hum Reprod 2013; 28: 2972–2980

van Loendersloot, L., Repping, S., Bossuyt, P.M., van der Veen, F., van Wely, M. **Prediction models in *in vitro* fertilization; where are we? A mini review.** J Adv Res 2014; 5: 295–301

Zhu, J., Lin, S., Li, M., Chen, L., Lian, Y., Liu, P., Qiao, J. **Effect of *in vitro* culture period on birthweight of singleton newborns.** Hum Reprod 2014; 29: 448–454

Received 5 March 2018; received in revised form 24 September 2018; accepted 25 September 2018.