Human Reproduction vol 11 no.10 pp 2298–2303, 1996

Prognostic models for the probability of achieving an ongoing pregnancy after in-vitro fertilization and the importance of testing their predictive value

A.M.Stolwijk[1,5], G.A.Zielhuis[1], C.J.C.M.Hamilton[2], H.Straatman[1], J.M.G.Hollanders[2], H.J.M.Goverde[2,3], P.A. van Dop[4] and A.L.M.Verbeek[1]

[1]Department of Medical Informatics, Epidemiology and Statistics, University of Nijmegen, P.O. Box 9101, NL-6500 HB Nijmegen, [2]Department of Obstetrics and Gynaecology, University Hospital Nijmegen, [3]Laboratory of Endocrinology and Reproduction, University Hospital Nijmegen, and [4]Department of Obstetrics and Gynaecology, Catharina Hospital, Eindhoven, The Netherlands

[5]To whom correspondence should be addressed

The aim of this study was to create reliable models to predict the probability of achieving an ongoing pregnancy during in-vitro fertilization (IVF) treatment: model A, at the start of the first treatment, model B, at the time of embryo transfer, and model C, during the second treatment at the end of the first IVF treatment. Prognostic models were created using data from the University Hospital Nijmegen (n = 757) and applied to the data from the Catharina Hospital Eindhoven (n = 432), The Netherlands, to test their predictive performance. The predictions of model B (made at time of embryo transfer) were fairly good (c = 0.672 in the test population). For instance, 93% of the patients who had a predicted probability of achieving an ongoing pregnancy of <10% did not achieve an ongoing pregnancy. However, the predictions of the other two models (A and C) for Eindhoven were less reliable. The predictive value of model C was fairly high in Nijmegen (c = 0.673). Its poor performance in the test population may be explained partly by differences in effectiveness of the ovulation stimulation protocols and the decision about when to discontinue the cycle. Thus, before using prognostic models at an IVF centre, their reliability at that specific centre should be tested.

Key words: in-vitro fertilization/ongoing pregnancy/pregnancy/prognosis/validation

Introduction

The probability that a patient will achieve an ongoing pregnancy should be evaluated as accurately as possible before a patient enters a programme for in-vitro fertilization (IVF) and during the course of her treatment with IVF. In addition to the age of the woman and the aetiology of infertility (the standard indicators for success), better rules would be welcome for physicians when counselling a patient. Potential predictors of IVF success are: patient characteristics at entry to the programme, characteristics of the treatment itself and during treatment, and intermediate results.

Most studies on factors that may predict pregnancy after treatment with IVF have investigated only a few indicators, for instance age and the type of infertility (Piette *et al.*, 1990; Hull *et al.*, 1992; Check *et al.*, 1993), baseline follicle stimulating hormone (FSH), luteinizing hormone (LH) and oestradiol concentrations (Padilla *et al.*, 1990), ovulation stimulation treatment and ovarian response (Dor *et al.*, 1992), endometrium thickness and uterine artery flow (Spernol *et al.*, 1993), sperm characteristics (Enginsu *et al.*, 1992) and age, oestradiol concentration, number and quality of oocytes and embryos (Fluker *et al.*, 1993). However, various simultaneous factors may influence the probability of achieving an ongoing pregnancy after IVF. It would therefore be desirable to create a model to predict the probability of achieving an ongoing pregnancy which includes all the relevant factors. Until now, only a few attempts have been made to do this for IVF (Hughes *et al.*, 1989; Haan *et al.*, 1991) and other assisted reproductive techniques (Guzick *et al.*, 1989; Nelson *et al.*, 1993). In the study by Hughes *et al.* (1989), age and failed fertilization due to poor sperm quality had a predictive value for success in subsequent IVF cycles. Haan *et al.* (1991) found that the probability of achieving an ongoing pregnancy after IVF treatment was increased by the presence of idiopathic infertility and decreased by the presence of a male factor, one ovary, the woman's age ≥ 36 years, primary infertility of at least 5 years duration and by a higher number of previous IVF treatments. Multivariate prognostic models should not be confused with explanatory models such as recently published by Roseboom *et al.* (1995). They discussed a multivariate model to explain the variation in the probability of pregnancy after embryo transfer. The variation was explained by the woman's age, average embryo morphology score, number of transferred embryos and an interaction term between tubal pathology and the woman's age. However, exclusion of the main effect of tubal pathology in the model makes a meaningful interpretation of the multivariate model difficult (Breslow and Day, 1980) and may cause bias (Kleinbaum *et al.*, 1982). Moreover, their statement in the results section '...with a 1 year increase of age, the probability of pregnancy for non-tubal patients decreased by 21%...' is obviously mistaken as a result of a wrong interpretation of the odds ratio in their study. Critical remarks can also be made about the methods used in the other four studies mentioned above (Guzick *et al.*, 1989; Hughes *et al.*, 1989; Haan *et al.*, 1991; Nelson *et al.*, 1993). All the cycles were combined, irrespective of the number of previous IVF treatments and the number of treatments per patient. Some studies based the inclusion of factors on statistical significance of the relationship in univariate analyses, which can be influenced by other factors, instead of on the increase in

the predictive power in multivariate models. Moreover, the predictions of these models were never tested in other populations. Thus, the validity of these prognostic models when used at other IVF centres can be questioned.

The purpose of this study is to create reliable models to predict the probability of achieving an ongoing pregnancy during the first or second treatment cycles with IVF. We used data from the University Hospital Nijmegen, The Netherlands, to develop the models, and data from another centre to test their predictive value.

Materials and methods

To develop the prognostic models, data were used from couples who were treated by IVF for the first time in the period March 1991 to January 1995 at the University Hospital, Nijmegen, The Netherlands. During this period IVF treatment hardly changed. To test these models, data were used from the Catharina Hospital, Eindhoven, The Netherlands. Guidelines on indications for IVF treatment in the Netherlands have been described by Jansen (1993). In short, couples are only offered IVF treatment in case of bilateral tubopathology, in cases of unilateral tubopathology, male factor, endometriosis or cervical factor when other infertility treatments had not resolved the problem, and in case of idiopathic infertility after an infertility duration of at least 3 years. For both populations data were only included if the complete IVF treatment had been carried out at that particular IVF centre, no donor oocytes had been used and no intracytoplasmatic sperm injection (ICSI) had been performed. Patient characteristics prior to treatment are given in Table I.

Ongoing pregnancy was defined as a pregnancy which continued for longer than 12 weeks after embryo transfer. To predict the probability of achieving an ongoing pregnancy, three models were developed that employed different moments of prediction. To predict the probability of achieving an ongoing pregnancy during the first IVF treatment, model A was made at the start and model B at the time of embryo transfer. For the prognosis of achieving an ongoing pregnancy during the second IVF treatment, model C was created at the end of the first IVF treatment. Table II presents the number of patients and pregnancies at each prediction moment.

Model A

This model was based on predictions made at the start of the first IVF cycle regarding the probability of achieving an ongoing pregnancy during the first IVF cycle. To develop this model, data were available from 757 couples whose first IVF cycle took place in Nijmegen. To induce ovulation, all the patients received a long protocol of gonadotrophin-releasing hormone (GnRH) agonist (usually Leuprolide; Abbott B.V., Amstelveen, The Netherlands or Suprefact; Hoechst Holland N.V., Amsterdam, The Netherlands) that was started on day 21 of the previous cycle and human menopausal gonadotrophin (HMG, Humegon; Organon Int. B.V., Oss, The Netherlands). Additionally, from August 1991 to January 1994, all the patients received oral contraceptives during the cycle that preceded the IVF cycle. To improve synchronization of follicle growth, some women received oral contraceptives before or after this period. To test the model, data were available from 432 couples from Eindhoven who underwent their first IVF treatment between January 1990 and June 1995 (another five couples were excluded from this population because information about the occurrence of an ongoing pregnancy was lacking). In this test population, the type of ovulation induction used most often (92.2%) was a short protocol of GnRH agonist (usually Suprefact; Hoechst Holland N.V.) and HMG (Humegon, Organon Int. B.V.), in a few cases supplemented by progestins in the preceding cycle.

Potential prognostic factors for the model that employed the onset of the first IVF cycle as the moment of prediction could only consist of information known at that moment, i.e. patient characteristics: age, period of infertility, reproductive history, basal FSH, indication(s) for IVF treatment, one or both ovaries present, sperm characteristics, anti-sperm antibodies in the woman or man, and information about the treatment protocol being used at that time: type of hormonal ovulation stimulation, maximum number of embryos that would be transferred, timing of human chorionic gonadotrophin (HCG) administration and type of culture medium. In Nijmegen, data on the duration of infertility were only available from patients who started IVF treatment between 1993–1994. Therefore the effect of the duration of infertility could only be estimated using the data from these 383 couples. Donor spermatozoa had not been used in Nijmegen, but it had been used in the test population in four and six patients during the first and second IVF cycles respectively. If donor spermatozoa had been used, the sperm characteristics were considered to be good and the indication for IVF 'male factor' was considered to be absent. We disregarded the results of cryopreserved embryo transfer.

Model B

Based on predictions made at the time of embryo transfer regarding the probability of achieving an ongoing pregnancy during the first IVF cycle. Only the data from couples who underwent embryo transfer during the first cycle were used to develop this model. Data were available from 604 (79.8% of the 757) couples from Nijmegen. To test the model, data could be used from 300 (69.4% of the 432) couples from Eindhoven. At this moment, information was added about preceding events during the cycle as potential prognostic factors, i.e. quality and number of oocytes retrieved, number of oocytes fertilized, quality and number of embryos transferred and whether the transfer had been uncomplicated as indicated by the use of a Wallace catheter, because in difficult cases a stiffer, Frydman catheter was used. In addition, information was known about the experience of the physician who performed the puncture and transfer; this could be used as a potential prognostic factor. Again, the results of cryopreserved embryo transfers were disregarded.

Model C

Based on predictions made at the end of the first IVF cycle regarding the probability of achieving an ongoing pregnancy during the second IVF cycle. To create this model, data were used from couples who did not have an ongoing pregnancy after the first IVF cycle or after a transfer of cryopreserved embryos and who started a second IVF cycle. In Nijmegen and in Eindhoven, 454 and 278 couples started a second IVF cycle respectively. In Eindhoven, information about ongoing pregnancy was lacking for three couples during the second cycle, so the data from 275 couples could be used for the test. In addition to the factors mentioned above, the pregnancy test result after the first IVF cycle was a potential prognostic factor in this model.

Statistical analysis

Models were developed by using logistic regression analysis. The first step was to develop a prognostic model based on patient characteristics and, if appropriate, the intermediate IVF treatment results. The second step was to evaluate whether treatment characteristics added any prognostic value to the model. The third step was to test the model.

Criteria for accepting variables as predictive factors in the model were based on statistical significance and added prognostic value, evaluated by using the c index [i.e. (number of concordant pairs +

Table I. Patient characteristics of the populations at the start of the first in-vitro fertilization (IVF) cycle

	Nijmegen (n = 757)					Eindhoven (n = 432)				
	Min.	Max	Mean	SD	Median	Min.	Max.	Mean	SD	Median
Woman's age (years)	22	47	32.9	4.0	33	21	43	31.8	4.1	32
Duration of infertility (years)[a]	0	20.5	4.4	2.8	4.0	0	20.5	3.7	2.7	3.5
Basal FSH (IU/l)	<0.6	23	6.1	2.8	5.7	NA				
	n		Percentage[b]			n		Percentage[c]		
≥1 Preceding gestations	256		33.8			162		37.5		
≥1 Preceding spontaneous abortions	126		16.6			29		8.3		
≥1 Preceding ectopic pregnancies	62		8.2			23		6.6		
≥1 Preceding deliveries	138		18.2			50		14.2		
Indication for IVF										
Tubal exclusively	168		22.2			147		34.6		
Tubal and other(s)	137		18.1			42		9.9		
Male factor exclusively	133		17.6			94[d]		22.1		
Male factor and other(s)	190		25.1			34[e]		8.0		
Endometriosis exclusively	44		5.8			34		8.0		
Endometriosis and other(s)	119		15.7			33		7.8		
Cervical factor exclusively	27		3.6			1		0.2		
Cervical factor and other(s)	91		12.0			1		0.2		
Idiopathic infertility	138		18.2			96		22.6		
Two ovaries	708		93.7			395		93.4		
Sperm characteristics										
≥20 × 10^6/ml	630		83.2			NA				
≥60% Normal forms	433		57.2			NA				
≥50% Motile	416		55.0			NA				
Quality of motility ≥4[f]	660		87.2			NA				
Anti-sperm antibodies, ♂ or ♀	66		8.7			10		2.3		
In sperm	38		5.0			NA				
In woman's serum	29		3.8			NA				
Use of donor spermatozoa	0		0.0			4		0.9		

NA = no information available
[a]Number of missing values for duration of infertility in Nijmegen n = 374.
[b]Number of missing values in Nijmegen: for two ovaries n = 1, anti-sperm antibodies ♂ or ♀, and in sperm n = 2
[c]Number of missing values in Eindhoven for ≥ 1 preceding spontaneous abortions, ectopic pregnancies, deliveries respectively n = 81, 81, 80, for the indications of IVF n = 7, for two ovaries n = 9
[d]Donor spermatozoa were used for three patients.
[e]Donor spermatozoa were used for one patient (the other indication for IVF was tubal factor)
[f]On a scale from 1 (worst) to 5 (best).

Table II. Number of patients and ongoing pregnancies

	Nijmegen			Eindhoven		
	No.	Pregnancies		No.	Pregnancies	
		n	%		n	%
At start of first IVF	757	88	11.6	432	46	10.6
At embryo transfer of first IVF	604	88	14.6	300	46	15.3
At start of second IVF	454	61	13.4	275	29	10.5

IVF = in-vitro fertilization.

0.5 × the number of tied pairs)/total number of pairs] (Harrell *et al.*, 1982; 1996). The c can be interpreted as the probability of a correct prediction for a random pair of a woman with an ongoing pregnancy and a woman without a pregnancy. It is equal to the area under a receiver operating characteristic (ROC) curve (Hanley and McNeil, 1982). For the development of a prognostic model, the erroneous exclusion of any prognostic factors (because of too little power) would be more deleterious than including too many factors. Therefore these criteria were given a high and low cut-off point respectively; P <0.10 and c >0.005. The variables were selected according to a method akin to a stepwise selection method. Here, the selection criteria is based not only on a P value (< 0.10), but also on a change in c (> 0.005). Special attention was given to multicollinearity. If this was present, only the variable with the highest predictive power was included in the multivariate model. If a variable did not meet the criteria in a univariate analysis, it thus could still be included in the prognostic model if the variable met the criteria when it was included in a multivariate model, i.e. after taking into account the prognostic value of other variables. In addition, a variable was omitted from the model if another factor was a stronger predictor and showed

Table III. Prognostic models for the probability of achieving an ongoing pregnancy (P) during the first in-vitro fertilization (IVF) or second IVF cycle

Model	Ln [P/(1-P)] =	SE(β)	$-2\ln(L_1/L_2)$ df	Nijmegen at development		Eindhoven at testing	
			P value	N^a	c	N^a	c
A	−0.3350	0.9503	14.04	757	0.612	431	0.497
	+0.8151 × ≥1 preceding gestation	0.2349	df = 2				
	−0.0620 × woman's age (years)	0.0297	P = 0.0009				
B	−4.2034	0.5399	48.96	603	0.721	171	0.672
	+0.5290 × ≥1 preceding gestation	0.2422	df = 4				
	+0.0630 × no. fertilized oocytes	0.0260	P = 0.0001				
	+0.3464 × no. transferred embryos	0.1711					
	+0.4377 × no. transferred embryos of at least good quality	0.1297					
C	−4.0236	0.7812	20.88	454	0.673	271	0.528
	+0.9886 × woman's age ≤30 years	0.4146	df = 4				
	+0.6001 × woman's age 31–35 years	0.3886	P = 0.0003				
	−0.8412 × idiopathic infertility	0.4537					
	+1.8638 × embryo transfer during first IVF cycle	0.7336					

[a]Patients with missing values on one or more of the variables were excluded, i.e. for Nijmegen 0, 1 and 0, and for Eindhoven 1, 129 and 4 for model A (prediction at start of first IVF cycle regarding probability during first cycle), model B (prediction at embryo transfer regarding probability during first cycle) and model C (prediction at end of first IVF cycle regarding probability during second cycle) respectively

no additional predictive value. For sperm characteristics combined variables were created and their predictive value was evaluated against that of the separate sperm characteristics

To test the predictive validity of the models, the data from the other centre were applied. As the data from Nijmegen contained more potential predictors than the data from Eindhoven, the models selected as the best predictive could not always be fully tested. If a specific variable was lacking, the model was modified, if possible, by exchanging it with a similar variable, or otherwise by excluding the variable. To evaluate the reliability of the model, the c was calculated. If the model had reasonable prognostic value, the predicted probability and the observed result of IVF were compared.

Results

The models for predicting the probability of achieving an ongoing pregnancy, developed with the data from Nijmegen and tested with the data from Eindhoven, are presented in Table III.

Model A

During the first IVF cycle, 88 (11.6%) out of the 757 women from Nijmegen and 46 (10.6%) out of the 432 women from Eindhoven achieved an ongoing pregnancy. The only factors that had predictive value were a previous gestation and the woman's age. During testing, this model did not show any predictive value when applied to the data from Eindhoven (c = 0.497).

Model B

Embryo transfer was performed in 604 (79.8%) out of the 757 couples from Nijmegen in the first IVF cycle. In Eindhoven, embryo transfer was performed in 300 (69.4%) out of the 432 couples. The ongoing pregnancy rate per transfer was 14.6% in Nijmegen and 15.3% in Eindhoven. The prognostic model included the factors: at least one preceding gestation, the number of fertilized oocytes, the number of transferred embryos and the number of transferred embryos of at least good quality. The probability of achieving an ongoing pregnancy increased if there had been a preceding gestation and the higher the numbers. During the test, this model showed good predictive value (c = 0.672) and good predictive performance, as shown in Table IV. For instance, 93% of the women with a predicted probability of achieving an ongoing pregnancy of <10% did not achieve an ongoing pregnancy after embryo transfer.

Model C

To predict the probability of achieving an ongoing pregnancy during the second IVF cycle, only the data from the couples who received a second treatment could be used. Of the 454 couples who received a second IVF treatment in Nijmegen, 61 (13.4%) achieved an ongoing pregnancy. In Eindhoven this occurred in 29 (10.5%) out of the 275 couples who underwent a second IVF treatment. The best prognostic model is shown in Table III. Of prognostic value were: the woman's age in age-groups, the presence of idiopathic infertility and embryo transfer during the first IVF cycle. However, this model did not show any predictive value in the test population (c = 0.528).

Discussion

This study showed that models for prediction of ongoing pregnancy due to IVF treatment can be developed with a fairly high prognostic value. However, this does not imply that the same models are predictive for patients treated at another clinic or even at the same clinic. Of the three models, only the one that made a prediction at the time of embryo transfer was fairly reliable in the other population. The other two models that made predictions at the start of treatment or after

Table IV. Predicted and observed percentages and numbers of women with an ongoing pregnancy during the first in-vitro fertilization (IVF) treatment at the time of embryo transfer in Eindhoven

Observed	Predicted probability (%)							
	0–<5	5–<10	10–<15	15–<20	20–<25	25–<30	≥30	Total
Percentage ongoing pregnancy	0	8	16	19	23	25	50	15
No women pregnant	0	5	6	6	3	2	3	25
Total number of women	16	60	37	31	13	8	6	171[a]

[a]No prediction could be made for 129 women, because no information was available about the number of transferred embryos of at least good quality.

the first IVF cycle, however, seemed to be of little value when used in Eindhoven. Although model B, which predicts at time of embryo transfer, is of little clinical importance, it gives information about the reasons for the inadequacy of the prediction at the start of the cycle For the two models at the start of the cycle, the ovarian response and oocyte aspiration are very important, but cannot be included as prognostic factors in the models because this information is not available at the start of the treatment, whereas in the model that made a prediction at time of embryo transfer, the number and quality of the retrieved oocytes are potential prognostic factors. Therefore, one explanation for the poor reliability might be differences in the effectiveness of the ovulation simulation protocols, the long protocol of GnRH agonist in Nijmegen and the short protocol in Eindhoven. No oocyte aspiration was performed during the first IVF treatment in 7.4% (56 out of the 757) and 21.8% (94 out of the 432) of the women from Nijmegen and Eindhoven, respectively. During the second IVF treatment, these percentages were 4.6% (21 out of the 454) and 15.3% (42 out of the 275) respectively. Not only might the effectiveness of the ovulation stimulation protocol have influenced the cancellation rate, but also the timing of this decision differed between the two centres. During the first IVF cycle, the percentage of cancelled cycles for the reason of too many follicles was only 1.8% in Nijmegen, but was as high as 31.5% in Eindhoven. This decision was made in Nijmegen if >25 follicles were present in combination with an oestradiol concentration of >20 000 pmol/l, whereas in Eindhoven, cycles were cancelled when >20 follicles were present. Whether the models developed in Nijmegen can make more accurate predictions if they are applied to an IVF centre that uses a long protocol of GnRH agonist and with fewer cancelled cycles remains to be seen.

The present models were adapted to make testing possible, given the information available in the test population. The changes were negligible. Models A and C were not changed at all. In model B the number of follicles >15 mm was initially included in the model, but because of lack of this information in the Eindhoven population, it was exchanged with the number of fertilized oocytes. Moreover, in model B the sperm characteristics <60% normal forms and/or $<20\times10^6$ spermatozoa per ml added minor predicting value, and were excluded from the model. Note that basal FSH had no additional predictive value, nor had the indications for IVF, except for idiopathic infertility in model C.

For prognosis, the predictive value of a positive test and of a negative test are of more practical value than the sensitivity and specificity of a test. The predictive value of a positive test is the proportion of patients with a positive test who achieve an ongoing pregnancy, and the predictive value of a negative test is the proportion of the patients with a negative test who do not achieve an ongoing pregnancy. Thus, they illustrate whether the prognosis was right, whereas the sensitivity and specificity of a test indicate whether the patients who achieved an ongoing pregnancy were classified well by the test. All these measures can be easily calculated using the data of Table IV. For instance, assume the cut-off point for the test to be a predicted probability of 5%; the test is positive if the predicted probability is ≥5% and negative if <5%. The positive predictive value of this test is 16% (25/155) and the negative predictive value is 100% (16/16). This demonstrates that the test can indicate patients who do not achieve an ongoing pregnancy after IVF, but cannot predict who achieves an ongoing pregnancy. The sensitivity and specificity of this test are 100% (25/25) and 11% (16/146) respectively.

Obviously, clinicians select their patients before treatment with IVF. If the study populations had included more extreme groups, those with a very high or a very low probability of success, then the reliability of the prognosis would have been better. The models we created only apply to populations that lie within the range of the characteristics presented in Table I. As women of 40 years of age or older were poorly represented in Nijmegen (n = 34), the models may not be valid for them. In addition, information on the duration of infertility was only available from 383 patients in Nijmegen The potential prognostic effect of the duration of infertility might not have been detected because of too few observations.

As the data were gathered retrospectively, it was not always possible to obtain full sets of information from the two databases. In some cases data were missing, or they were not present in the desired form. Moreover, the two hospitals had their own method of performing IVF and the patient populations might have differed on other aspects than those studied Therefore it was more difficult to create a model that would make reliable predictions than if the data had been gathered in a standardized way for the purpose of prognostic studies at hospitals which use the same treatment protocols and the same definitions for each variable. To make it possible to create reliable prognostic models, we recommend setting up uniform national registries which also contain information about the basic fertility workup.

The importance of testing prognostic models is evident. Untested prognostic models can be worthless when used for prediction at another (or possibly even the same) IVF clinic.

Before a model can be used by another IVF centre, it should be tested with retrospective data from that centre, to establish whether it is a predictive model in that centre. Even before a model is implemented in the centre where it was developed, it should be tested with an entirely separate set of data from the same centre before one can rely on its predictive properties.

Acknowledgements

This study was made possible by the data from the IVF centres at the University Hospital Nijmegen and Catharina Hospital Eindhoven.

References

Breslow, N.E. and Day, N.E (1980) *Statistical Methods in Cancer Research. Volume 1 The Analysis of Case-Control Studies* International Agency for Research on Cancer, Lyon, p 197

Check, J H., Lurie, D , Dietterich, C *et al* (1993) Adverse effect of a homogeneous hyperechogenic endometrial sonographic pattern, despite adequate endometrial thickness on pregnancy rates following in-vitro fertilization. *Hum. Reprod.*, **8,** 1293–1296

Dor, J , Seidman, D S , Ben Shlomo, I *et al.* (1992) The prognostic importance of the number of oocytes retrieved and estradiol levels in poor and normal responders in *in vitro* fertilization (IVF) treatment *J Assist Reprod. Genet* , **9,** 228–232

Enginsu, M E , Pieters, M.H , Dumoulin, J C *et al.* (1992) Male factor as determinant of in-vitro fertilization outcome *Hum. Reprod.*, **7,** 1136–1140

Fluker, M.R., Siu, C.K , Gunby, J *et al.* (1993) Cycle characteristics and outcome in relation to ovarian response during *in vitro* fertilization *J Assist Reprod Genet* , **10,** 504–512

Guzick, D S , Balmaceda, J P , Ord, T *et al.* (1989) The importance of egg and sperm factors in predicting the likelihood of pregnancy from gamete intrafallopian transfer *Fertil Steril* , **52,** 795–800

Haan, G , Bernardus, R E , Hollanders, J.M G *et al.* (1991) Results of IVF from a prospective multicentre study *Hum. Reprod.*, **6,** 805–810

Hanley, J A and McNeil, B.J (1982) The meaning and use of the area under a Receiver Operating Characteristic (ROC) curve *Radiology,* **143,** 29–36

Harrell, F E , Califf, R M , Pryor, D B *et al.* (1982) Evaluating the yield of medical tests *J Am. Med. Assoc* , **247,** 2543–2646.

Harrell, F E , Lee, K.L. and Mark, D.B. (1996) Multivariable prognostic models issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors *Stat Med.*, **15,** 361–387

Hughes, E G , King, C and Wood, E C (1989) A prospective study of prognostic factors in *in vitro* fertilization and embryo transfer *Fertil. Steril.*, **51,** 838–844

Hull, M G R , Eddowes, H.A , Fahy, U *et al.* (1992) Expectations of assisted conception for infertility *Br. Med. J* , **304,** 1465–1469

Jansen, C.A M (1993) Indicaties voor *in vitro* fertilisatie *Ned. Tijdschr Obstet Gynaecol.*, **106,** 35–36

Kleinbaum, D.G , Kupper, L L and Morgenstern, H (1982) *Epidemiologic Research. Principles and Quantitative Methods* Van Nostrand Reinhold Company, New York, p 454

Nelson, J R., Huppert, L., Corson, S L *et al* (1993) Predicting success of gamete intrafallopian transfer *Fertil. Steril.*, **60,** 116–122.

Padilla, S L , Bayati, J and Garcia, J E (1990) Prognostic value of the early serum estradiol response to leuprolide acetate in *in vitro* fertilization. *Fertil Steril.*, **53,** 288–294.

Piette, C , De Mouzon, J , Bachelot, A *et al.* (1990) In-vitro fertilization influence of women's age on pregnancy rates *Hum. Reprod.*, **5,** 56–59

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

Spernol, R , Hecher, K., Schwarzgruber, J *et al* (1993) Doppler-Flow-Messungen in der Arteria uterina Ein Prognosefaktor für den Erfolg bei der Behandlung durch IVF? *Ultraschall Med.*, **14,** 175–177

Received on March 18, 1996, accepted on August 8, 1996