Human Reproduction, Vol.31, No.9 pp. 1942–1951, 2016
Advanced Access publication on July 12, 2016 doi:10.1093/humrep/dew146

human reproduction

ORIGINAL ARTICLE *Andrology*

Prediction model for live birth in ICSI using testicular extracted sperm

A.M. Meijerink[1,†*], M. Cissen[2,†], M.H. Mochtar[3], K. Fleischer[1], I. Thoonen[1], A.A. de Melker[3], A. Meissner[3,4], S. Repping[3], D.D.M. Braat[1], M. van Wely[3], and L. Ramos[1]

[1]Department of Obstetrics and Gynaecology, Division of Reproductive Medicine, Radboud University Medical Center, PO Box 9101, 6500 HB Nijmegen, The Netherlands [2]Department of Obstetrics and Gynaecology, Jeroen Bosch Hospital, PO Box 90153, 5200 ME 's-Hertogenbosch, The Netherlands [3]Center for Reproductive Medicine, Department of Obstetrics and Gynaecology, Academic Medical Center, PO Box 22660, 1100 DE Amsterdam, The Netherlands [4]Department of Urology, Academic Medical Center, PO Box 22660, 1100 DE Amsterdam, The Netherlands

*Correspondence address. Department of Obstetrics and Gynaecology (791), Radboud University Medical Center, PO Box 9101, 6500 HB Nijmegen, The Netherlands. E-mail: aukje.meijerink@radboudumc.nl

Submitted on March 13, 2016; resubmitted on May 6, 2016; accepted on May 26, 2016

STUDY QUESTION: Which parameters have a predictive value for live birth in couples undergoing ICSI after successful testicular sperm extraction (TESE-ICSI)?

SUMMARY ANSWER: Female age, a first or subsequent started TESE-ICSI cycle, male LH, male testosterone, motility of the spermatozoa during the ICSI procedure and the initial male diagnosis before performing TESE were identified as relevant and independent parameters for live birth after TESE-ICSI.

WHAT IS KNOWN ALREADY: In reproductive medicine prediction models are used frequently to predict treatment success, but no prediction model currently exists for live birth after TESE-ICSI.

STUDY DESIGN, SIZE, DURATION: A retrospective cohort study between 2007 and 2015 in two academic hospitals including 1559 TESE-ICSI cycles. The prediction model was developed using data from one centre and validation was performed with data from the second centre.

PARTICIPANTS/MATERIALS, SETTING, METHODS: We included couples undergoing ICSI treatment with surgically retrieved sperm from the testis for the first time. In the development set we included 526 couples undergoing 1006 TESE-ICSI cycles. In the validation set we included 289 couples undergoing 553 TESE-ICSI cycles. Multivariable logistic regression models were constructed in a stepwise fashion ($P < 0.2$ for entry). The external validation was based on discrimination and calibration.

MAIN RESULTS AND THE ROLE OF CHANCE: We included 224 couples (22.3%) with a live birth in the development set. The occurrence of a live birth was associated with lower female age, first TESE-ICSI cycle, lower male LH, higher male testosterone, the use of motile spermatozoa for ICSI and having obstructive azoospermia as an initial suspected diagnosis. The area under the receiver operating characteristic (ROC) curve was 0.62. From validation data, the model had moderate discriminative capacity (*c*-statistic 0.67, 95% confidence interval: 0.62–0.72) but calibrated well, with a range from 0.06 to 0.56 in calculated probabilities.

LIMITATIONS, REASONS FOR CAUTION: We had a lack of data about the motility of spermatozoa during TESE, therefore, we used motility of the spermatozoa used for ICSI after freeze-thawing, information which is only available during treatment. We had to exclude data on paternal BMI in the model because too many missing values in the validation data hindered testing. We did not include a histologic diagnosis, which would have made our data set less heterogeneous and, finally, our model may not be applicable in centres which have a different policy for the indication for performing sperm extraction. The prognostic value of the model is limited because of a low 'area under the curve'.

WIDER IMPLICATIONS OF THE FINDINGS: This model enables the differentiation between couples with a low or high chance to reach a live birth using TESE-ICSI. As such it can aid in the counselling of patients and in clinical decision-making.

STUDY FUNDING/COMPETING INTEREST(S): This study was partly supported by an unconditional grant from Merck Serono (to D.D.M.B. and K.F.) and by the Department of Obstetrics and Gynaecology of Radboud University Medical Center, Nijmegen, The Netherlands, the Department of Obstetrics and Gynaecology, Jeroen Bosch Hospital, Den Bosch, The Netherlands, and the Department of Obstetrics and

[†] The authors consider that the first two authors should be regarded as joint First Authors.

Gynaecology, Academic Medical Center, Amsterdam, The Netherlands. Merck Serono had no influence in concept, design, nor elaboration of this study.

TRIAL REGISTRATION NUMBER: Not applicable.

Key words: prediction model / sperm injection / ICSI / sperm retrieval / pregnancy / azoospermia / testicular sperm extraction

Introduction

One revolutionary achievement in assisted reproductive techniques (ART) has been the introduction of ICSI in 1992 (Palermo *et al.*, 1992). ICSI allowed couples with severe oligozoospermia to father their genetically own offspring. Subsequently, ICSI was used in conjunction with testicular sperm extraction (TESE) in men with azoospermia (Devroey *et al.*, 1995).

Spermatozoa are found in ~50% of the men with non-obstructive azoospermia (NOA) who undergo TESE (Tournaye *et al.*, 1995; Chan and Schlegel, 2000). Surgical sperm retrieval carries a small but existing complication risk, such as loss of significant amounts of testicular tissue, haematoma, inflammatory changes and permanent devascularisation (Schlegel and Su, 1997). Only one out of seven men undergoing TESE eventually fathers a genetically own child (Vloeberghs *et al.*, 2015). We recently developed and externally validated a prediction model for obtaining spermatozoa with TESE in men with NOA (Cissen *et al.*, 2016).

Once spermatozoa are successfully retrieved, it is possible to start ICSI treatment. Although TESE-ICSI is a blessing for the NOA couple, it is in our opinion still important to inform patients of their realistic chances to conceive. Couples should have the possibility to consider whether it is worth the intense treatment burden that entails ovarian stimulation, oocyte retrieval, and not to forget the emotional impact of a failed treatment (Verhaak *et al.*, 2007; Rockliff *et al.*, 2014). For this we need a clinical prediction model which can predict the chance of a live birth adjusted for couple's characteristics.

In reproductive medicine several prediction models have been developed for spontaneous pregnancy, pregnancy after intrauterine insemination and pregnancy after IVF/ICSI (Leushuis *et al.*, 2009). In studies that predict the outcome of ICSI, it was demonstrated that maternal age is the most important predictive parameter next to the duration of infertility and obstetric history (Stolwijk *et al.*, 2000; Lintsen *et al.*, 2007). Nevertheless, it is unknown if these parameters play the same predictive role for couples who require TESE-ICSI, as these couples have no natural chances for pregnancy. It is possible that in these specific couples, other parameters will have an impact on the probability of getting pregnant.

The aim of this study was to develop and validate a model for couples who had a successful TESE, to predict their chance of having a live birth with TESE-ICSI.

Materials and Methods

Study design

Between 1 September 2007 and 1 May 2014 we performed a retrospective cohort study, among couples undergoing TESE-ICSI at the Radboud university medical center, The Netherlands (Radboudumc). The data we collected from Radboudumc were used to develop a model to predict live birth after a TESE-ICSI cycle (development set). An ICSI cycle was defined as a fresh cycle and the corresponding cryo embryo cycle(s) derived from it; each cycle was considered as a separate unit of analysis. Between 1 August 2007 and 1 September 2015, we collected data for validation of the model in couples undergoing TESE-ICSI in the Academic Medical Center (AMC), The Netherlands (validation set).

Ethical approval

Until 2014 in the Netherlands, a TESE procedure was allowed only in research settings. The protocol for this study was approved by the Dutch Central Committee on Research involving Human Subjects (NL12408.000.06 CCMO, The Hague, The Netherlands). All couples signed an informed consent for treatment and follow-up before participating in this study.

Study population

The study population contained couples eligible for ICSI treatment. We included men with azoospermia—defined as no spermatozoa found in the sediment of a centrifuged sample, confirmed in at least two semen analyses (WHO, 2009)—who, after a complete andrologic evaluation by an urologist, were diagnosed with NOA and subsequently underwent a successful TESE procedure or men who were initially suspected for obstructive azoospermia (OA), underwent an unsuccessful percutaneous epididymal sperm aspiration (PESA) procedure, but a successful TESE procedure, i.e. viable spermatozoa were found. The TESE procedure was performed in both clinics. We excluded men with deletions in the *AZF-a* or *AZF-b* region of the Y chromosome and Klinefelter syndrome.

TESE procedure

All TESE procedures were performed by a trained urologist. The description of the local clinical protocol has been published (Hessel *et al.*, 2013). In summary, a conventional longitudinal testicular biopsy according to the method described by Silber was performed in all men (Silber, 2010) before the start of female hormonal stimulation, to make sure that viable sperm were available at oocyte retrieval. If spermatozoa were found in the testicular biopsy, aliquots of sperm cell suspensions were cryopreserved and female hormonal stimulation was started and the sperm cell suspension was thawed at oocyte retrieval. In case no viable sperm was found, a second TESE was performed with the agreement of the patient, and in the absence of physical or medical contraindication, and fresh sperm cells were used at oocyte retrieval. Surplus aliquots of spermatozoa cell suspensions of the second TESE were cryopreserved as well.

ICSI cycles

All couples underwent at least one ICSI cycles. Controlled ovarian hyperstimulation was performed with recombinant FSH (Puregon or Gonal-F) or human urinary FSH (Menopur or Fostimon) after pituitary down-regulation with a GnRH agonist or antagonist. Oocyte retrieval took place by ultrasound-guided needle aspiration, 34–36 h after hCG (5000 or 10 000 IU) administration. Only metaphase II, morphologically normal oocytes were injected with frozen–thawed or fresh spermatozoa. In case of immotile spermatozoa, viable spermatozoa were selected using the tail touch procedure (de Oliveira *et al.*, 2004; Hessel *et al.*, 2015). The tail touch procedure involves the recording of sperm tail flexibility after 2 h of incubation. The tail of the spermatozoon is

gently touched with an ICSI micropipette, and when the tail is flexible the spermatozoon is considered viable.

Embryo scoring and evaluation for transfer was based on fragmentation and the number of blastomeres, assessed by a qualified embryologist or laboratory technician using a light inverted microscope, at standard set points: fertilization, early cleavage stage, Day 2 and Day 3 of development (ESHRE, 2000).

Intrauterine embryo transfer was performed at the third, fourth or fifth day after oocyte retrieval. Single embryo transfer (SET) or double embryo transfer (DET) was performed depending on female age, female medical history, the couple's preference and national policy. National embryo transfer policy changed over time, which brought about compulsory SET in women <38 years in the first two cycles for the last 3 years of the study period. No embryo transfer was performed in case of absence of fertilization, abnormal embryos or in case of (risk for) ovarian hyperstimulation syndrome (OHSS). Luteal phase was supported by vaginal administration of 600 mg progesterone per day. An ongoing pregnancy was defined as the appearance of a fetal heartbeat examined by ultrasound after 12 weeks of gestational age. Live birth was defined as the complete expulsion or extraction from its mother of a product of fertilization, irrespective of the duration of the pregnancy, which, after such separation, breathes or shows any other evidence of life such as heart beat, umbilical cord pulsation, or definite movement of voluntary muscles, irrespective of whether the umbilical cord has been cut or the placenta is attached (Zegers-Hochschild et al., 2009).

Building the predictive model

We used live birth as the primary end-point of the study. We defined one delivery as the birth of one singleton baby or multiples. We developed a model by calculating the probability of live birth after a TESE-ICSI cycle. We considered each started cycle including corresponding cryo embryo cycle(s) as a separate unit of analysis. Model building was based on TESE-ICSI data from the development set (Radboudumc).

Based on the current literature, previous prediction models and our own expectations, we identified a number of candidate predictors (Friedler et al., 2002; Leushuis et al., 2009; Boitrelle et al., 2011; Tehraninejad et al., 2012; Ramasamy et al., 2013; van Loendersloot et al., 2013). The candidate baseline parameters/covariates were as follows:

Type of infertility (primary/secondary);
Duration of infertility (months);
Female age (years);
Parity (*n*);
Average menstrual cycle length (days);
Uterine abnormalities (yes/no);
Antral follicle count before stimulation (number of follicles <11 mm);
Alcohol use (self-reported; yes/no) for male and female;
Smoking status (self-reported; yes/no) for male and female;
BMI at baseline (kg/m^2) for male and female;
Male age (years);
Male testosterone (nmol/l);
Male inhibin B (ng/l);
Male FSH (IU/l);
Male LH (IU/l);
Total testicular volume (cc);
Suspected primarily diagnosis of azoospermia (OA/NOA) before sperm retrieval.

NOA was defined as azoospermia, in combination with either small testes (volume per testis <15 ml), elevated level of FSH (>10 IU/l) and/or decreased level of inhibin B (<150 ng/l) (Adamopoulos and Koukkou, 2010; Jungwirth et al., 2012) and without evidence of obstruction. In men with evidence of obstruction and with normal testes volume, FSH and inhibin B levels, the initial suspected diagnosis was OA.

The candidate cycle parameters/covariates were:

Number of TESE-ICSI cycles;
Spermatozoa (fresh or frozen–thawed);
Motility of spermatozoa (oocytes injected with motile spermatozoa/immotile spermatozoa or a combination of both for each individual cycle);
Number of oocytes retrieved.

Ideally, all parameters in a prognostic model are available before start of the treatment. In this case it would be preferable to include the motility of spermatozoa found at the TESE procedure. However, due to a lack of these data we decided to use the motility of the spermatozoa used for ICSI (in most cases after freeze-thawing). Besides motility, three other cycle depended parameters were also analysed as described above.

Statistical analysis

For each candidate prognostic variable, the association with occurrence of an ongoing pregnancy leading to a live birth was assessed using the χ score test in a logistic regression model. Colinearity between variables was assessed to prevent the inclusion of redundant variables in the model. After the inclusion of female age, covariates were selected using forward selection ($P < 0.20$ for entry). Backward elimination ($P > 0.20$ for removal) confirmed the covariate selection for the final model. All subjects were included in the final models with missing covariate values imputed using linear regression. First-order interaction terms and quadratic terms were tested, but not found to be statistically significant.

We first analysed our data with generalized estimating equations (GEE) and afterwards with logistic regression. The point estimates and confidence intervals (CI) after analysis with GEE were almost identical to those of logistic regression. As logistic models are easier to interpret and the point estimates did not differ, we decide to use multivariable logistic regression to develop a model.

For the final logistic regression model the receiver operating characteristic (ROC) curve was plotted and the area under the curve (AUC, or *c*-statistic) was calculated.

As the capacity of a variable to predict live birth may vary in a series of ICSI cycles, we explicitly tested statistically for interactions between included predictors and ICSI cycle number. In deciding between competing expressions of related parameters, we used Akaike's information criterion in variable selection.

These characteristics are data driven and presumably too optimistic. Optimism-corrected values were calculated using leave-one-out cross-validation, i.e. the regression coefficients associated with the 'final model' were re-estimated with each subject left out in turn. We then combined the 'leave-one-out' regression coefficient with the subject's covariate values in order to mimic the prediction of the outcome for each subject. Finally, a logistic regression model was fitted with the resulting 'leave-one-out' prognostic index as the only covariate in order to obtain the optimism-corrected AUC. A histogram (not shown) displaying the distribution of the predicted probabilities was plotted.

Model validation

A crucial aspect of a prediction model derived from one data set is the wider applicability to a data set from another centre or from a different time period. The idea of validating a prognostic model is generally taken to mean establishing that it works satisfactorily for patients other than those from whose data the model was derived (Altman and Royston, 2000). External model validation was based on the TESE-ICSI data from the validation set (at AMC) and focused on two aspects: discrimination and calibration (Leushuis et al., 2009).

Discrimination was measured by the area under the ROC curve, the *c*-statistic. This statistic ranges from 0.5 (no discrimination) to 1 (perfect

discrimination). Calibration refers to correspondence between the predicted probabilities and the observed proportions. Calibration was assessed visually by comparing predicted probabilities and observed proportions after dividing patients in six groups based on their predicted probability and, more formally, by fitting a logistic regression model. All analyses were performed using IBM SPSS Statistics 22 (Chicago, IL, USA) and STATA 14 (Texas, USA).

Results

In this study we included in total 815 couples who had undergone 1559 TESE-ICSI cycles for the development and validation set. Figure 1 shows the steps to take for a couple undergoing TESE-ICSI treatment and Table I shows the baseline characteristics. For the development set,

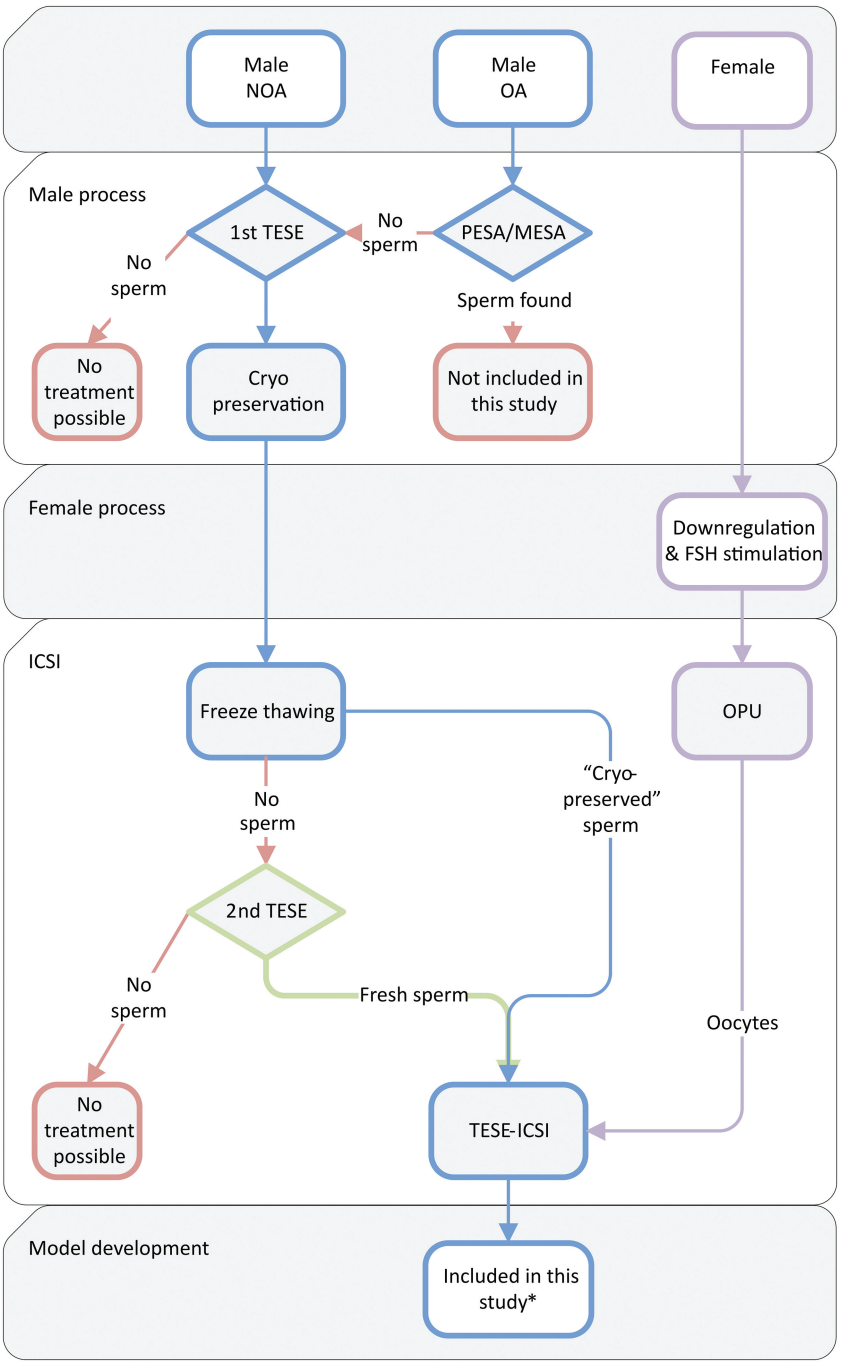

Figure 1 Steps in the testicular sperm extraction (TESE)-ICSI treatment. Suspected diagnosis based on medical history, physical examination and hormonal values. NOA, non-obstructive azoospermia; OA, obstructive azoospermia; TESE, testicular sperm extraction; PESA, percutaneous epididymal sperm aspiration; MESA, microsurgical epididymal sperm aspiration; OPU, oocyte pick-up (retrieval). *Following cycles of couples who already had a previous live birth after TESE-ICSI during the study period were excluded.

Table I Baseline characteristics of the testicular sperm extraction (TESE)-ICSI cycles included in this study.

	Total *n* = 1559	Development set *n* = 1006	Validation set *n* = 553
No. of couples (*n*)	815	526	289
Clinical characteristics, mean (±SD)			
Female age years	32.8 (4.5)	32.4 (4.4)	33.7 (4.5)
Male age years	38.5 (8.5)	38.2 (8.1)	39.1 (9.0)
Type of infertility per cycle			
Primary infertility, *n* (%)		830 (82.5)	
Secondary infertility, *n* (%)		176 (17.5)	
Duration of infertility months, mean (±SD)		41.5 (25.2)	
Female parameters per cycle, *n* (%)			
Endometriosis		18 (1.8)	
Polycystic ovary syndrome		50 (5.0)	
Congenital uterine anomaly		4 (4.0)	
Acquired uterine anomaly		13 (12.9)	
Male hormones at baseline, mean (±SD)			
Male testosterone, nmol/l	15.2 (5.8)	14.8 (6.0)	16.0 (5.4)
Male inhibin B, ng/l		92.2 (79)	
Male FSH, IU/l	13.9 (11.6)	13.6 (11.3)	15.0 (12.8)
Male LH, IU/l	6.4 (4.1)	6.0 (3.6)	7.5 (5.0)
No. of ICSI cycles, *n* (%)			
1st cycle	815 (52.3)	526 (52.3)	289 (52.3)
2nd cycle	466 (29.9)	306 (30.4)	160 (28.9)
3rd cycle	222 (14.2)	137 (13.6)	85 (15.4)
≥4th cycle	56 (3.6)	37 (3.7)	19 (3.4)
Sperm characteristics			
Sperm condition in ICSI cycle, *n* (%)			
Fresh	239 (15.3)	200 (19.9)	39 (7.1)
Frozen thawed	1317 (84.5)	806 (80.1)	511 (92.4)
Unknown	3 (0.2)	0	3 (0.5)
Sperm motility in ICSI cycle, *n* (%)			
Motile	773 (49.6)	641 (63.7)	132 (15.7)
Immotile	180 (11.5)	138 (13.7)	42 (7.6)
Both motile and immotile	305 (19.6)	218 (21.7)	87 (23.9)
Motility unknown	301 (19.3)	9 (0.9)	292 (52.8)
Laboratory data per cycle, mean (±SD)			
No. of oocytes	10.5 (5.6)	10.3 (5.2)	10.9 (6.2)
No. of oocytes fertilized		4.4 (3.1)	
No. of 2PN embryos	4.1 (3.2)	4.3 (3.0)	3.8 (3.4)
No. of frozen embryos		0.6 (1.3)	
No. of embryos transferred	1.3 (0.7)	1.4 (0.6)	1.2 (0.8)
Pregnancies, *n* (%)			
No. of ongoing pregnancies	348 (22.3)	229 (22.8)	119 (21.5)
No. of live births	337 (21.6)	224 (22.3)	113 (20.4)

No.: number; Acquired uterine anomaly: uterus myomatosus or endometrial polyp; PN: pronuclei.

we included 526 couples who had undergone 1006 TESE-ICSI cycles of which 379 men (72.1%) were diagnosed with NOA and 147 men (27.9%) with suspected OA but with no sperm retrieval after PESA or microsurgical epididymal sperm aspiration (MESA) and successful testicular sperm retrieval (OA>NOA diagnosis group). In 73 couples (80 cycles) there was no embryo transfer: 66 cycles resulted in a total

Table II Multivariable analysis for predicting live birth after a TESE-ICSI cycle.

Predictors	OR	95% CI	*P*-value	AUC	AUC corrected
Female age^2	0.99	0.98–0.99	0.02	0.56	0.56
No. of ICSI cycles	0.76	0.56–1.04	0.07	0.58	0.58
Male LH	0.94	0.98–0.99	0.02	0.62	0.62
Male testosterone	1.03	1.00–1.05	0.06	0.61	0.61
Sperm motility	1.82	1.09–3.04	0.02	0.60	0.60
NOA versus OA>NOA	1.31	0.90–1.97	0.09	0.63	0.63

Suspected diagnosis before sperm retrieval based on medical history, physical examination and hormonal values.
OA>NOA: suspected OA but with no sperm retrieval after epididymal sperm aspiration.
OR: odds ratio; CI: confidence interval; AUC: area under the curve; Female age^2: age square; NOA: non-obstructive azoospermia; OA: obstructive azoospermia.

failure of fertilization, in 10 cycles there was no suitable embryo for transfer present and in 4 cycles a fresh embryo transfer was withheld because of imminent OHSS. The cumulative ongoing pregnancy rate per cycle was 229 (22.8%) of which 224 (22.3%) resulted in a live birth of 1 or more children: 196 singletons, 27 twins and 1 triplet. Three pregnancies were terminated because of congenital anomalies: 1 with both a hygroma colli and an omphalocele, 1 with trisomy 18 and 1 because of craniocephalic malformation. Two pregnancies involved a stillbirth at 20 and 41 weeks of gestational age, respectively.

For the external validation set, we included 289 couples who had undergone 553 TESE-ICSI cycles of which 169 men (58.4%) were diagnosed with NOA and 105 men (36.3%) were in the OA>NOA diagnosis group. In 15 men (5.2%) the initial diagnosis was unclear due to missing hormonal values. There were 119 ongoing pregnancies (21.5%) of which 113 (20.4%) resulted in a live birth of 1 or more children.

Univariable analysis showed that younger women, younger men, men with lower LH, lower FSH and higher testosterone levels, and those with motile spermatozoa used for ICSI and having an OA as initial diagnosis and those undergoing the first ICSI cycle, had significantly higher chances of a live birth. None of the hormone values were below the limit of detection.

We included six predictors in the final multivariable logistic regression model: female age, cycle number, male LH level, male testosterone level, sperm motility and suspected diagnosis before sperm retrieval (OA versus NOA). When we added a square term for female age, we did not find interactions, nor did interaction terms improve the model.

In the development set three variables we had selected for the final model, were incomplete, i.e. male LH (9%), male testosterone (9%) and sperm motility (1%) had missing values. In the validation set four variables we had selected were incomplete, i.e. male LH (9%), male testosterone (3%), sperm motility (53%) and initial diagnosis (4%) had missing values. Whether or not data were missing was not associated with the occurrence of an ongoing pregnancy leading to a live birth and were not significant in the analysis described above.

The multivariable analysis is presented in Table II. We did not find a significant additional effect of ICSI cycle number, nor did we find any significant interactions between the identified predictors and cycle number. For this reason, we used the same point estimates for all predictors and we included cycle number as a predictor.

The calculated probabilities of a live birth for the 526 couples in the development set ranged from 0.03 to 0.47, with a mean of 0.19. Twenty-five per cent of the TESE-ICSI cycles had a probability of a pregnancy less than 0.14, 25% had a probability between 0.14 and 0.19, 25% a probability between 0.19 and 0.24, and 25% had a probability exceeding 0.24.

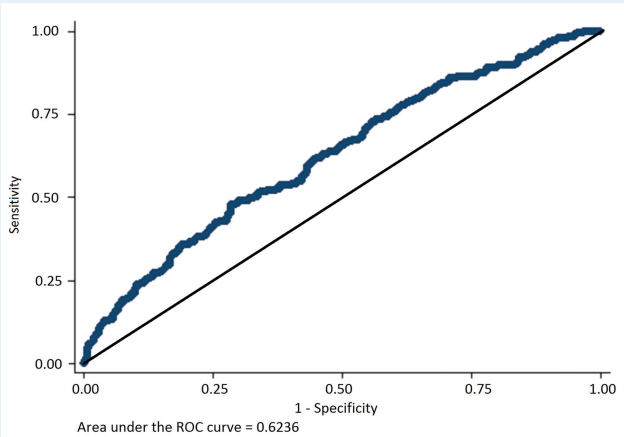

Figure 2 Receiver operating characteristic curve (ROC) of the multivariable logistic regression model for the prediction of live birth after TESE-ICSI (model development).

In the development set the model had moderate discriminative capacity with a *c*-statistic of 0.63 (95% CI: 0.59–0.67) and in the over optimism corrected model 0.62 (95% CI: 0.58–0.68) (Fig. 2). The model calibrated well; the goodness-of-fit test (Hosmer–Lemeshow) showed no significant miscalibration ($P = 0.79$). Figure 3A shows the calibration plot in the development set. In the case of perfect calibration, all points would be on the diagonal, the line of equality and average probabilities correspond perfectly to the observed live birth rates. Our calibration plot showed that the model calibrated well. In the calibration model in the development set, the estimated intercept was 0.02 (95% CI: −0.05 to 0.09) and the slope 1.07 (95% CI: 0.89–1.25). The intercept approached zero and the slope unity.

External validation

External model validation was based on the TESE data from the AMC in Amsterdam (validation set) and focused on two aspects: discrimination and calibration (Leushuis *et al.*, 2009).

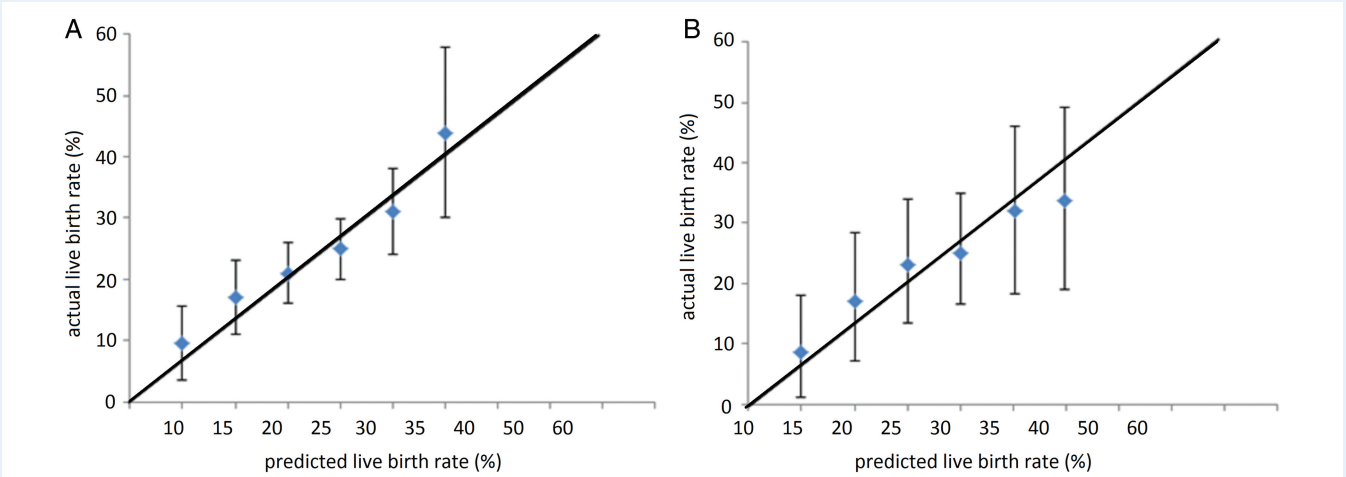

Figure 3 Calibration plots with calculated probability on the *X*-axis and observed proportion on *Y*-axis for the development set (**A**) and validation set (**B**).

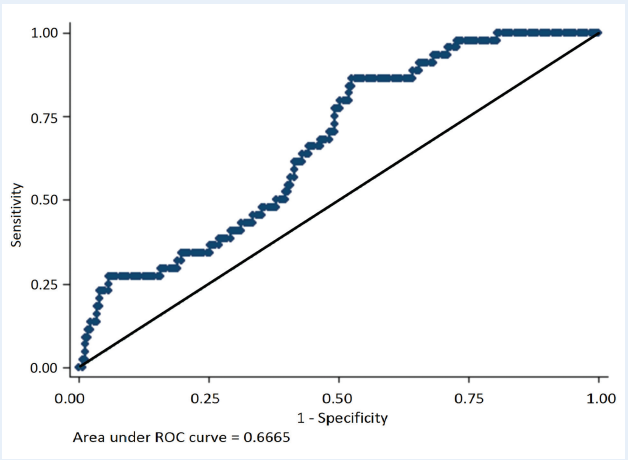

Figure 4 Receiver operating characteristic curve of the multivariable logistic regression model for the prediction of live birth after TESE-ICSI (model validation).

Discrimination is the ability of the model to distinguish between cases with and without the event of interest, in this case between men with successful sperm retrieval with TESE and men where no spermatozoa could be found. Discrimination was measured by the area under the ROC curve, i.e. *c*-statistic. This statistic ranges from 0.5 (no discrimination) to 1 (perfect discrimination). Calibration refers to correspondence between the predicted probabilities and the observed probabilities. In the external validation set the model had slightly better discriminative capacity. The *c*-statistic was 0.67 (95% CI: 0.62–0.72) (Fig. 4). The model calibrated well; the goodness-of-fit test (Hosmer–Lemeshow) showed no significant miscalibration ($P = 0.73$). The calibration plot visually resembles the calibration of the development set though with larger 95% boundaries and a range from 0.06 to 0.56 in calculated probabilities (Fig. 3B).

Table III presents the prediction model and shows an example of calculated probabilities of a live birth in a TESE-ICSI cycle for five hypothetical couples based on our prediction model. It should be taken into account whether or not the spermatozoa used for ICSI are motile. For example couple C of which the female age is 34 years, who will have their first ICSI cycle, with a male LH of 4 IU/l and a male testosterone of 20 nmol/l in a man diagnosed with NOA, the chance of a live birth in the first ICSI cycle is 22% when there are enough motile spermatozoa for the injection of all oocytes and 13% when there are not.

Validation of the model without sperm motility

The sperm motility was only known for 261 (47%) cycles of the external validation set therefore we also made a model without sperm motility. This model, including all previously selected variables while excluding sperm motility, had a *c*-statistic of 0.61 (95% CI: 0.57–0.65) in the development set and 0.62 (95% CI: 0.55–0.69) in the validation set. However this model did not calibrate well and therefore this model is not discussed further.

Discussion

In our analysis of 526 infertile couples suffering from azoospermia who had 1006 TESE-ICSI cycles, we found that a live birth was associated with a lower female age, the first versus subsequent TESE-ICSI cycles, lower male LH, higher male testosterone, availability of motile spermatozoa for ICSI and having OA as an initial diagnosis. After model development using multivariable logistic regression analysis, we found an AUC of 0.62, and good calibration. We validated our model externally with 553 cycles in 289 infertile couples in a second academic hospital performing TESE-ICSI during the same study period. In the external validation set the model had a slightly better discriminative capacity (AUC 0.67) and calibrated well.

We found similar predictors, such as female age and the number of started cycles (first versus subsequent), as previous studies on IVF/ICSI had found (Lintsen *et al.*, 2007; van Loendersloot *et al.*, 2013). These models are however not applicable to our couples since the females are not *per se* infertile and the azoospermic men carry specific predictors, not to be found in men with oligo or normospermia.

Table III Five hypothetical couples with the calculated probability of a live birth in a TESE-ICSI cycle.

	Couple A		Couple B		Couple C		Couple D		Couple E	
Female age (years)	24		30		34		38		40	
ICSI cycle	1		2		1		2		1	
Male LH (E/l)	4		8		4		10		6	
Male testosterone (nmol/l)	20		14		20		10		14	
NOA(=0) or OA>NOA (=1)	OA>NOA		NOA		NOA		NOA		OA>NOA	
Motile spermatozoa for ICSI, Yes (=1) or No (=0)	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Calculated probability of a live birth per cycle	0.31	0.20	0.16	0.09	0.22	0.13	0.06	0.04	0.10	0.06

Probability of live birth after a TESE-ICSI cycle = EXP (−7.659 + 0.442 × Female age − 0.008 × Female age^2 − 0.278 × No. ICSI cycle + 0.608 × Sperm motility + 0.026 × Testosterone − 0.06 × LH + 0.269 × (N)OA)/(1 + EXP (−7.659 + 0.442 × Female age − 0.008 × Female age^2 − 0.278 × No. ICSI cycle + 0.608 × Sperm motility + 0.026 × Testosterone − 0.06 × LH + 0.269 × (N)OA)).
Suspected diagnosis before sperm retrieval based on medical history, physical examination and hormonal values.
OA>NOA: suspected OA but with no sperm retrieval after epididymal sperm aspiration.

We found that a lower male LH and a higher male testosterone level are of predictive value for live birth after TESE-ICSI in our study group. Since the production of testosterone in the Leydig cells is regulated by LH, a lower male LH will indicate a (relative) normal testicular function because of negative feedback by testosterone on LH. In men with NOA the hypothalamic-pituitary axis is often already deregulated leading to low testosterone levels, but obesity might also contribute to low testosterone. Unfortunately, we could not include paternal BMI as a predictor in our model, because of too many missing values. Perhaps, if we had complete data, paternal BMI might play a minor predictive role. One previous study found that sperm retrieval rates were similar in normal weight and obese men, but a lower male BMI predicted a higher chance of clinical pregnancy after TESE-ICSI (Ramasamy *et al.*, 2013). This study however did not report on LH and testosterone. Our findings support the hypothesis that deregulation of the hypothalamic–pituitary–gonadal axis may influence the quality of sperm in terms of the ability to conceive. Future research should explore whether or not paternal obesity and perhaps coinciding low testosterone levels effects live birth rates. Since in the majority of the cases the genetic cause of the azoospermia is still unknown (Ezeh, 2000), it is possible that genetic effects may also influence the live birth rate.

In our study sperm motility was found as an independent predictor for live birth. In ejaculated sperm, motility is a marker for sperm integrity and is associated with better chromatin condensation and less DNA-damage (Ramos and Wetzels, 2001; Moskovtsev *et al.*, 2009; Ortega *et al.*, 2011). It is unknown whether motility is also a marker for integrity of testicular sperm, since testicular sperm is not meant to be motile and chromatin condensation is incomplete.

The strength of our study is that we used data from the only two centres in the Netherlands who were allowed to perform TESE-ICSI until May 2014. Before the study period TESE-ICSI was not allowed due to restrictions in government policies based on concerns about the safety of the procedure for the men and health of the children born from the use of testicular sperm. It enabled us to develop and also externally validate a prediction model with live birth as primary outcome.

The applicability of our results in different clinical settings is subject to certain limitations. For instance, in this study we included all couples with azoospermic men, in which TESE was the only suitable method for sperm retrieval and who underwent TESE-ICSI, i.e. we did not differentiate between patients diagnosed with NOA and patients who were suspected of OA but with a previous failed PESA/MESA. We did not include a histologic diagnosis, which would have made our data set less heterogeneous. For those couples with an OA>NOA diagnosis (suspected OA but with no sperm retrieval after epididymal sperm aspiration), the sperm retrieval rate after TESE is higher. Our model may not be applicable in centres which have a different policy for the indication of performing PESA and TESE. Moreover, in these cycles the chance to find enough motile spermatozoa for injection of all oocytes, which we found as an independent predictive factor for live birth, is also higher. Finally, the prognostic value of the model is limited because of a low 'area under the curve'.

In fertility treatment for couples with men diagnosed with azoospermia there are several steps to take, e.g. TESE, oocyte retrieval, successful thawing of TESE spermatozoa, injection of oocytes, fertilization, embryo transfer, pregnancy and finally giving birth. A clear starting point should be determined for developing a prediction model for these couples. In this study we choose this point after successful TESE; we included couples when the oocyte retrieval resulted in the injection of oocytes with testicular extracted spermatozoa. We have also developed a model which predicts the chance of finding spermatozoa in men with NOA (Cissen *et al.*, 2016). The predictive capacity our model was fair with an AUC of 0.69 in the development set and 0.65 in the validation set. The calibration indicated that our model could distinguish men with a poor prognosis from men with a good prognosis.

Ideally, all parameters in a prognostic model are known before start of the treatment. However, we had a lack of data of the motility of spermatozoa at TESE. Therefore, we decided to use the motility of the spermatozoa used for ICSI. In the counselling of couples one should discuss both options, i.e. with or without motile spermatozoa used for ICSI. The motility of spermatozoa at TESE (before freeze thawing) may give an indication for the chance of having motile spermatozoa for ICSI. Sperm cells extracted after a first TESE, and remaining after a first use for TESE-ICSI, could be used in a second TESE-ICSI treatment. Our model provides some insight and may help in the counselling of couples about their chances of success during the fertility treatment

and might help them in their decision whether or not to continue treatment after a failed cycle.

Without TESE-ICSI NOA couples do not have a chance to conceive. Physicians can involve patients in the process of making decisions about their health so that patients receive care that meets their needs and wishes. This is called 'shared decision-making' (Legare *et al.*, 2010). Using shared decisions, the physician and the couple can assess the burden of the treatment and the chance of a live birth (Baysal *et al.*, 2015). Several studies found that patients who are involved in decision-making are less prone to experience decisional conflict and regret (Legare *et al.*, 2010; Bastings *et al.*, 2014). Using our model a physician is better able to inform the couple about their actual chances.

In conclusion, in this study we developed and externally validated a prediction model which is able for the first time to predict the chance of live birth after TESE-ICSI. Using this model it is possible to counsel couples individually about their chances to conceive and becoming parents. Considering their individual chance to conceive and weighing their personal options will help couples in deciding to start or continue treatment or not. Our study revealed that besides the known factors, such as female age and the number of started cycles, a number of male-related factors (i.e. male LH, male testosterone, motility of spermatozoa used for ICSI and the initial male diagnosis) were also predictors for live birth after TESE-ICSI.

Acknowledgements

We would like to thank Dr K.W. D'Hauwers for her help in data collection, and performing TESE procedures in Radboud university medical center, The Netherlands.

Authors' roles

All authors qualify for authorship by contributing substantially to this manuscript. All authors developed the original concept of this study collectively. Data collection was performed by A.M., I.T, and M.C., statistical analysis by M.v.W. All authors have contributed to critical discussion and reviewed the final version of the manuscript and approve it for publication.

Funding

This study was funded by the Department of Obstetrics and Gynaecology, Radboud university medical center, Nijmegen, The Netherlands and partly supported by unrestricted grants from Merck Serono (the Netherlands) to D.D.M.B. and K.F. However, Merck Serono did not have any influence in concept and design or elaboration of this study. Further funding was received from the Department of Obstetrics and Gynaecology, Jeroen Bosch Hospital, Den Bosch, The Netherlands, and from the Department of Obstetrics and Gynaecology, Academic Medical Center, Amsterdam, The Netherlands.

Conflict of interest

This study was partly funded by unrestricted grants to D.D.M.B. and K.F.; however, Merck Serono had no influence in concept, design, nor elaboration of this study.

References

Adamopoulos DA, Koukkou EG. 'Value of FSH and inhibin-B measurements in the diagnosis of azoospermia'—a clinician's overview. *Int J Androl* 2010; **33**:e109–e113.

Altman DG, Royston P. What do we mean by validating a prognostic model? *Stat Med* 2000;**19**:453–473.

Bastings L, Baysal O, Beerendonk CC, IntHout J, Traas MA, Verhaak CM, Braat DD, Nelen WL. Deciding about fertility preservation after specialist counselling. *Hum Reprod* 2014;**29**:1721–1729.

Baysal O, Bastings L, Beerendonk CC, Postma SA, IntHout J, Verhaak CM, Braat DD, Nelen WL. Decision-making in female fertility preservation is balancing the expected burden of fertility preservation treatment and the wish to conceive. *Hum Reprod* 2015;**30**:1625–1634.

Boitrelle F, Robin G, Marcelli F, Albert M, Leroy-Martin B, Dewailly D, Rigot JM, Mitchell V. A predictive score for testicular sperm extraction quality and surgical ICSI outcome in non-obstructive azoospermia: a retrospective study. *Hum Reprod* 2011;**26**:3215–3221.

Chan PT, Schlegel PN. Nonobstructive azoospermia. *Curr Opin Urol* 2000; **10**:617–624.

Cissen M, Meijerink AM, d'Hauwers KW, Meissner A, van der Weide N, Mochtar MH, de Melker AA, Ramos L, Repping S, Braat DDM *et al.* Prediction model for obtaining spermatozoa with TESE in men with non-obstructive azoospermia. *Hum Reprod* 2016; doi:10.1093/humrep/dew147.

de Oliveira NM, Vaca Sanchez R, Rodriguez Fiesta S, Lopez Salgado T, Rodriguez R, Bethencourt JC, Blanes Zamora R. Pregnancy with frozen-thawed and fresh testicular biopsy after motile and immotile sperm microinjection, using the mechanical touch technique to assess viability. *Hum Reprod* 2004;**19**:262–265.

Devroey P, Liu J, Nagy Z, Goossens A, Tournaye H, Camus M, Van Steirteghem A, Silber S. Pregnancies after testicular sperm extraction and intracytoplasmic sperm injection in non-obstructive azoospermia. *Hum Reprod* 1995;**10**:1457–1460.

ESHRE Atlas of Embryology. *Hum Reprod* 2000;**15**.

Ezeh UI. Beyond the clinical classification of azoospermia: opinion. *Hum Reprod* 2000;**15**:2356–2359.

Friedler S, Raziel A, Strassburger D, Schachter M, Soffer Y, Ron-El R. Factors influencing the outcome of ICSI in patients with obstructive and non-obstructive azoospermia: a comparative study. *Hum Reprod* 2002; **17**:3114–3121.

Hessel M, Ramos L, Hulsbergen AF, D'Hauwers KW, Braat DD, Hulsbergen-van de Kaa CA. A novel cell-processing method 'AgarCytos' in conjunction with OCT3/4 and PLAP to detect intratubular germ cell neoplasia in non-obstructive azoospermia using remnants of testicular sperm extraction specimens. *Hum Reprod* 2013;**28**:2608–2620.

Hessel M, Robben JC, D'Hauwers KW, Braat DD, Ramos L. The influence of sperm motility and cryopreservation on the treatment outcome after intracytoplasmic sperm injection following testicular sperm extraction. *Acta Obstet Gynecol Scand* 2015;**94**:1313–1321.

Jungwirth A, Giwercman A, Tournaye H, Diemer T, Kopa Z, Dohle G, Krausz C; European Association of Urology Working Group on Male Infertility. European Association of Urology guidelines on Male Infertility: the 2012 update. *Eur Urol* 2012;**62**:324–332.

Legare F, Ratte S, Stacey D, Kryworuchko J, Gravel K, Graham ID, Turcotte S. Interventions for improving the adoption of shared decision making by healthcare professionals. *Cochrane Database Syst Rev* 2010; CD006732.

Leushuis E, van der Steeg JW, Steures P, Bossuyt PM, Eijkemans MJ, van der Veen F, Mol BW, Hompes PG. Prediction models in reproductive medicine: a critical appraisal. *Hum Reprod Update* 2009;**15**:537–552.

Lintsen AM, Eijkemans MJ, Hunault CC, Bouwmans CA, Hakkaart L, Habbema JD, Braat DD. Predicting ongoing pregnancy chances after IVF and ICSI: a national prospective study. *Hum Reprod* 2007;**22**:2455–2462.

Moskovtsev SI, Willis J, White J, Mullen JB. Sperm DNA damage: correlation to severity of semen abnormalities. *Urology* 2009;**74**:789–793.

Ortega C, Verheyen G, Raick D, Camus M, Devroey P, Tournaye H. Absolute asthenozoospermia and ICSI: what are the options? *Hum Reprod Update* 2011;**17**:684–692.

Palermo G, Joris H, Devroey P, Van Steirteghem AC. Pregnancies after intracytoplasmic injection of single spermatozoon into an oocyte. *Lancet* 1992;**340**:17–18.

Ramasamy R, Bryson C, Reifsnyder JE, Neri Q, Palermo GD, Schlegel PN. Overweight men with nonobstructive azoospermia have worse pregnancy outcomes after microdissection testicular sperm extraction. *Fertil Steril* 2013;**99**:372–376.

Ramos L, Wetzels AM. Low rates of DNA fragmentation in selected motile human spermatozoa assessed by the TUNEL assay. *Hum Reprod* 2001; **16**:1703–1707.

Rockliff HE, Lightman SL, Rhidian E, Buchanan H, Gordon U, Vedhara K. A systematic review of psychosocial factors associated with emotional adjustment in in vitro fertilization patients. *Hum Reprod Update* 2014; **20**:594–613.

Schlegel PN, Su LM. Physiological consequences of testicular sperm extraction. *Hum Reprod* 1997;**12**:1688–1692.

Silber SJ. Sperm retrieval for azoospermia and intracytoplasmic sperm injection success rates—a personal overview. *Hum Fertil* 2010;**13**:247–256.

Stolwijk AM, Wetzels AM, Braat DD. Cumulative probability of achieving an ongoing pregnancy after in-vitro fertilization and intracytoplasmic sperm injection according to a woman's age, subfertility diagnosis and primary or secondary subfertility. *Hum Reprod* 2000;**15**:203–209.

Tehraninejad ES, Pourmatroud E, Sadighi Gilani MA, Rakebi M, Azimi Neko Z, Arabipoor A. Comparison of intracytoplasmic sperm injection outcomes between oligozoospermic, obstructive azoospermic and non-obstructive azoospermic patients. *Int J Fertil Steril* 2012;**6**:13–18.

Tournaye H, Camus M, Goossens A, Liu J, Nagy P, Silber S, Van Steirteghem AC, Devroey P. Recent concepts in the management of infertility because of non-obstructive azoospermia. *Hum Reprod* 1995;**10** (Suppl 1):115–119.

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

Verhaak CM, Smeenk JM, Evers AW, Kremer JA, Kraaimaat FW, Braat DD. Women's emotional adjustment to IVF: a systematic review of 25 years of research. *Hum Reprod Update* 2007;**13**:27–36.

Vloeberghs V, Verheyen G, Haentjens P, Goossens A, Polyzos NP, Tournaye H. How successful is TESE-ICSI in couples with non-obstructive azoospermia? *Hum Reprod* 2015;**30**:1790–1796.

WHO. WHO Laboratory Manual for the Examination of Human Semen and Sperm Cervical Mucus Interaction, 5th edn. Cambridge University Press, 2009.

Zegers-Hochschild F, Adamson GD, de Mouzon J, Ishihara O, Mansour R, Nygren K, Sullivan E, van der Poel S; International Committee for Monitoring Assisted Reproductive Technology; World Health Organization. The International Committee for Monitoring Assisted Reproductive Technology (ICMART) and the World Health Organization (WHO) Revised Glossary on ART Terminology, 2009. *Hum Reprod* 2009;**24**:2683–2687.