Human Reproduction, Vol.31, No.9 pp. 1934–1941, 2016
Advanced Access publication on July 12, 2016 doi:10.1093/humrep/dew147

human reproduction
ORIGINAL ARTICLE ***Andrology***

Prediction model for obtaining spermatozoa with testicular sperm extraction in men with non-obstructive azoospermia

M. Cissen[1,2,†], A.M. Meijerink[3,†], K.W. D'Hauwers[4], A. Meissner[2,5], N. van der Weide[2], M.H. Mochtar[2], A.A. de Melker[2], L. Ramos[3], S. Repping[2], D.D.M. Braat[3], K. Fleischer[3], and M. van Wely[2,*]

[1]Department of Obstetrics and Gynaecology, Jeroen Bosch Hospital, PO Box 90153, 5200 ME 's-Hertogenbosch, The Netherlands [2]Department of Obstetrics and Gynaecologie, Center for Reproductive Medicine, Academic Medical Center, PO Box 22660, 1100 DE Amsterdam, The Netherlands [3]Department of Obstetrics and Gynaecology, Division of Reproductive Medicine, Radboud University Medical Center, PO Box 9101, 6500 HB Nijmegen, The Netherlands [4]Department of Urology, Radboud University Medical Center, PO Box 9101, 6500 HB Nijmegen, The Netherlands [5]Department of Urology, Academic Medical Center, PO Box 22660, 1100 DE Amsterdam, The Netherlands

*Correspondence address. E-mail: m.vanwely@amc.uva.nl

Submitted on March 13, 2016; resubmitted on March 13, 2016; accepted on May 2, 2016

STUDY QUESTION: Can an externally validated model, based on biological variables, be developed to predict successful sperm retrieval with testicular sperm extraction (TESE) in men with non-obstructive azoospermia (NOA) using a large nationwide cohort?

SUMMARY ANSWER: Our prediction model including six variables was able to make a good distinction between men with a good chance and men with a poor chance of obtaining spermatozoa with TESE.

WHAT IS KNOWN ALREADY: Using ICSI in combination with TESE even men suffering from NOA are able to father their own biological child. Only in approximately half of the patients with NOA can testicular sperm be retrieved successfully. The few models that have been developed to predict the chance of obtaining spermatozoa with TESE were based on small datasets and none of them have been validated externally.

STUDY DESIGN, SIZE, DURATION: We performed a retrospective nationwide cohort study. Data from 1371 TESE procedures were collected between June 2007 and June 2015 in the two fertility centres.

PARTICIPANTS/MATERIALS, SETTING, METHODS: All men with NOA undergoing their first TESE procedure as part of a fertility treatment were included. The primary end-point was the presence of one or more spermatozoa (regardless of their motility) in the testicular biopsies.

We constructed a model for the prediction of successful sperm retrieval, using univariable and multivariable binary logistic regression analysis and the dataset from one centre. This model was then validated using the dataset from the other centre. The area under the receiver-operating characteristic curve (AUC) was calculated and model calibration was assessed.

MAIN RESULTS AND THE ROLE OF CHANCE: There were 599 (43.7%) successful sperm retrievals after a first TESE procedure. The prediction model, built after multivariable logistic regression analysis, demonstrated that higher male age, higher levels of serum testosterone and lower levels of FSH and LH were predictive for successful sperm retrieval. Diagnosis of idiopathic NOA and the presence of an azoospermia factor c gene deletion were predictive for unsuccessful sperm retrieval. The AUC was 0.69 (95% confidence interval (CI): 0.66–0.72). The difference between the mean observed chance and the mean predicted chance was $<$2.0% in all groups, indicating good calibration. In validation, the model had moderate discriminative capacity (AUC 0.65, 95% CI: 0.62–0.72) and moderate calibration: the predicted probability never differed by more than 9.2% of the mean observed probability.

LIMITATIONS, REASONS FOR CAUTION: The percentage of men with Klinefelter syndrome among men diagnosed with NOA is expected to be higher than in our study population, which is a potential selection bias. The ability of the sperm retrieved to fertilize an oocyte and produce a live birth was not tested.

[†] M.C. and A.M.M. consider that the first two authors should be regarded as joint First Authors.

WIDER IMPLICATIONS OF THE FINDINGS: This model can help in clinical decision-making in men with NOA by reliably predicting the chance of obtaining spermatozoa with TESE.

STUDY FUNDING/COMPETING INTEREST: 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 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 injections / ICSI / sperm retrieval / azoospermia

Introduction

Azoospermia is defined by the complete absence of spermatozoa in the ejaculate (WHO, 2009). It affects ~1% of the male population (Jarow et al., 1989) and can be divided into two main categories; mechanical obstruction along the seminal tract (obstructive azoospermia (OA)) and an intrinsic testicular impairment of sperm production (non-obstructive azoospermia (NOA)). NOA is diagnosed in 60% of azoospermic men (Jarow et al., 1989; Matsumiya et al., 1994). Using ICSI in combination with testicular sperm extraction (TESE) even men suffering from NOA are able to father their own biological child.

TESE is an invasive procedure with an, albeit small, risk of complications, including haematoma, inflammation and permanent devascularization, all possibly resulting in the loss of a significant amount of testicular tissue (Schlegel and Su, 1997; Donoso et al., 2007). Only in approximately half of the patients with NOA can testicular sperm be retrieved successfully (Chan and Schlegel, 2000; Colpi et al., 2005; Tournaye, 2010).

A number of factors have been suggested to be of predictive value to distinguish patients with a good chance to retrieve sperm cells from patients who have a poor chance. Factors such as testicular volume, serum FSH levels and serum inhibin B levels (Chen et al., 2010; Toulis et al., 2010; Bryson et al., 2014; Yildirim et al., 2014; Yang et al., 2015) have been used as variables in several prediction models (Samli and Dogan, 2004; Tsujimura et al., 2004; Boitrelle et al., 2011; Ramasamy et al., 2013). Unfortunately, all these models are of limited value. Most included only a limited number of patients and the models were not validated externally. Since a model tends to perform better in the population in which it has been constructed, external validation is a crucial step before the model can be used in daily practice across different centres (Bleeker et al., 2003).

It is evident that patients should be well informed about their chances of a successful TESE procedure and their likelihood to retrieve sperm cells before consenting to the procedure. It would therefore be of great value to be able to estimate an individual's chance of sperm retrieval to empower patients in their decision-making. The aim of this study is to develop an externally validated model to predict successful sperm retrieval with TESE in men with NOA in a large nationwide cohort.

Materials and Methods

Study design

We performed a nationwide retrospective cohort study among all men diagnosed with NOA undergoing a first TESE as part of a fertility treatment in the Radboud University Medical Center, Nijmegen (Radboudumc—development set) and the Academic Medical Center, Amsterdam (AMC—validation set), The Netherlands. Data were collected between 1 June 2007 and 1 June 2015. These two centres were the only centres offering TESE–ICSI during the study period until 1 October 2014, as legal restrictions in the Netherlands limited TESE–ICSI to be conducted only in research settings.

Ethical approval

The protocol for this multicentre study was approved by the Dutch Central Committee on Research involving Human Subjects (NL12408.000.06 CCMO, The Hague, The Netherlands). All couples signed informed consent for treatment and follow-up before participating in this study.

Study population

We included all men diagnosed with NOA and undergoing their first TESE as a part of fertility treatment. NOA was defined as azoospermia without evidence of obstruction and with an elevated level of FSH (>10 UI/l), decreased level of inhibin B (<150 ng/l) and/or small testicular volume (<15 cc per testis) (Adamopoulos and Koukkou, 2010; Jungwirth et al., 2015). Before surgical sperm extraction, each male underwent a complete andrological evaluation by a urologist. We defined idiopathic NOA when there was no abnormality found in diagnostics such as Klinefelter syndrome, azoospermia factor (*AZF*)-deletions, cryptorchidism, infection and malignancy. Men with deletions of the *AZFa* or *AZFb* region of the Y chromosome were excluded for TESE in the Radboudumc. Until September 2010, men diagnosed with Klinefelter syndrome (47XXY) were excluded in both centres due to restrictions in government policies.

TESE procedure

A TESE procedure was performed under local anaesthesia or occasionally under general anaesthesia. A scrotal skin incision was made over the largest testicle or, in case of equal volume, the testicle with the better consistency. Thereafter, the tunica vaginalis was opened and, if necessary, the testis luxated outside the scrotum. The tunica albuginea was longitudinally incised. Either a longitudinal biopsy over the whole length (in case of uniform morphology of the seminiferous tubules) or targeted biopsies of the thicker tubules (in case of differences in quality of the tubules) were taken and immediately transported to the fertility laboratory. The biopsy was then subjected to mechanical dissection and cells present in the lumen of the tubules were extracted (Silber, 2000). The obtained cell suspension was directly examined for the presence of spermatozoa. If no spermatozoa were present in the initial biopsy, then a subsequent biopsy was taken from the contralateral testis. If spermatozoa were present, their number and motility were noted, and the cell suspension was cryopreserved, as described previously (Hessel et al., 2013).

Outcome

The presence of one or more spermatozoa (regardless of their motility) in the testicular biopsy was considered as successful sperm retrieval and was used as the primary end-point of the study.

Model building

A model was developed to calculate the probability of obtaining sperm with a TESE procedure. Data from Radboudumc were used for building this model (development set). Only the first performed TESE procedure for each couple was included for the analysis. Candidate prognostic parameters or covariates were male age (years), BMI (kg/m^2), smoking behaviour (self-reported; yes/no), alcohol consumption (self-reported; yes/no), duration of infertility (months), serum testosterone (nmol/l), inhibin B (ng/l), FSH (IU/l) and LH (IU/l) levels, total testicular volume (cc) (measured by physical examination) and aetiology of NOA (Klinefelter syndrome/*AZFc* deletion/cryptorchidism and/or orchidopexy/idiopathic/others).

For each candidate prognosticator, the association with successful sperm retrieval was assessed using the χ^2 test in a logistic regression model. Collinearity between variables was assessed to prevent the inclusion of redundant variables in the model. All cases were included in the final model, and cases with missing covariate values were imputed using multiple imputation. Missing data varied from 0% to ~6% for the potential predictor variables and these values were interpreted as missing at random. We checked the linearity of the association between the continuous variables and a successful sperm retrieval using cubic spline analyses and used transformation in case of non-linearity.

Statistical analysis

Covariates were selected using forward selection ($P < 0.15$ for entry). Backward elimination ($P > 0.15$ for removal) confirmed the covariate selection for the final model. First-order interaction terms and quadratic terms were tested, but not found to be significant.

For the final logistic regression model, we used the Akaike Information Criterion for each imputation set separately to account for differences between sets. Predictors for the final multivariable model were selected using the majority method. The receiver-operating characteristic (ROC) curve was plotted, and the area under the curve (AUC, or *c*-statistic) was calculated. These characteristics are data driven and presumably too optimistic, therefore the calculated values were denoted as 'apparent' AUC. Optimism-corrected values were calculated using leave-one-out cross-validation, i.e. regression coefficients associated with the 'final model' were re-estimated with each case left out in turn. We then combined the 'leave-one-out' regression coefficient with the case's covariate values in order to mimic the prediction of the outcome for each case. Finally, a logistic regression model was fitted with the resulting 'leave-one-out' prognostic index (PI) as the only covariate in order to obtain the optimism-corrected AUC. A histogram displaying the distribution of the predicted probabilities was plotted. A score chart (Hunault *et al.*, 2004) was constructed for easier application of the model.

Model 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).

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. Calibration was assessed visually by comparing predicted probabilities and observed probabilities after dividing men into four groups based on their predicted probability and, more formally, by fitting a logistic regression model with a single covariate for the so-called PI, a linear combination of the case's covariate values and the associated regression coefficients.

For all analyses, we used IBM SPSS Statistics 22 (Chicago, IL, USA) and STATA 14 (Stata-Corp, College Station, TX, USA).

Results

A total of 1371 first TESE procedures were included: 918 in the development set and 453 in the validation set. Baseline characteristics are shown in Table I. The sperm retrieval rate was 45.6% (419/918) in the development set and was 39.7% (180/453) in the validation set.

Model building

The spline function demonstrated a non-linear association between the continuous variable LH and successful sperm retrieval. We therefore transformed LH to better fit the data using a polynominal: LH + LH^2.

Univariable analysis confirmed that older men, men with higher testosterone and high inhibin B levels, men with a history of cryptorchidism and/or orchidopexy and men with a larger testicular volume had significantly higher chances of successful sperm retrieval. Diagnosis of idiopathic NOA, Klinefelter syndrome, detection of an *AZFc* deletion and high levels of FSH and LH were significantly associated with lower chances of successful sperm retrieval.

The multivariable logistic regression model (Table II) included male age, levels of testosterone, LH and FSH, diagnosis of idiopathic NOA and the presence of an *AZFc* deletion as independent predictors.

Cases with missing values for the covariates selected in the final model were evaluated separately. In these 81 cases (FSH ($n = 8$), LH ($n = 46$), testosterone ($n = 43$), *AZF* deletion ($n = 10$), idiopathic NOA ($n = 10$)) the fact of whether data were missing or was not associated with a successful sperm retrieval.

The calculated probability of sperm retrieval in the development set had a range from 6 to 93%, with a mean of 44% (Fig. 1A).

The model had moderate discriminative capacity in the development set. The *c*-statistic was 0.70 (95% confidence interval (CI): 0.66–0.73) and 0.69 (95% CI: 0.66–0.72) in the optimism-corrected model (Fig. 2A). In the calibration model in the development set, the estimated intercept was −0.01 (95% CI: −0.10 to 0.08) and the slope 1.03 (95% CI: 0.84–1.22). The intercept approached zero and the slope unity. The predicted probability of successful sperm retrieval was compared with the observed sperm retrieval rate in that category. The difference between the mean observed chance and the mean predicted chance was <2.0% in all groups, which indicates a good calibration of the prediction model in the development set (Table III and Fig. 3A).

The mathematical formulation for predicting the probability of sperm retrieval for an individual man is as follows: probability $= 1/[1 + \exp(-\beta)]$, where $\beta = -1.009 + (\text{male age} \times 0.058) + (\text{LH} \times -0.115) + (\text{LH}^2 \times 0.001) + (\text{FSH} - 0.019) + (\text{testosterone} \times 0.034) + (\textit{AZFc}\ \text{deletion} - 1.480) + (\text{idiopathic NOA} - 0.855)$. Table IV shows an example of calculated probabilities of successful sperm retrieval with TESE for five hypothetical men based on our prediction model.

External validation

The calculated probabilities of successful sperm retrieval in the validation set ranged from 4 to 92%, with a mean of 41%, indicative of a population with comparable probabilities to the men in the development set (Fig. 1B). The discriminative capacity was in the same range as that in the development set, with a *c*-statistic of 0.65 (95% CI: 0.59–0.70)

Table I Baseline characteristics of men with NOA who underwent a first TESE.

		Total (n = 1371)	Development set (n = 918)	Validation set (n = 453)
Clinical characteristics				
Age (years, SD)		34.29 (6.3)	34.05 (6.02)	34.78 (6.83)
Duration infertility (months, SD)		–	33.36 (25.24)	–
Male BMI (kg/m^2, SD)		–	26.04 (4.13)	–
Smoke *n* (%)	Yes	–	224 (24.6)	–
	No	–	688 (75.4)	–
Alcohol *n* (%)	Yes	–	705 (77.5)	–
	No	–	205 (22.5)	–
Testosterone (nmol/l, SD)		13.97 (5.46)	13.43 (5.33)	15.00 (5.55)
LH (IU/l,SD)		8.97 (4.99)	8.57 (4.67)	9.85 (5.53)
FSH (IU/l, SD)		22.13 (12.22)	21.57 (11.83)	23.27 (12.91)
Inhibin B (ng/l, SD)		39.04 (38.85)	40.09 (38.70)	35.89 (39.21)
Total testicular volume (cc, SD)		–	24.94 (9.16)	–
Diagnosis *n* (%)				
Idiopathic		531 (38.7)	341 (37.1)	190 (41.9)
Klinefelter		85 (6.2)	42 (4.6)	43 (9.5)
AFZc deletion		63 (4.6)	35 (3.8)	28 (6.2)
Genetic (others)		12 (0.9)	8 (0.9)	4 (0.9)
Cryptorchidism and/or orchidopexy		450 (32.8)	340 (37.0)	110 (24.3)
Others		209 (15.2)	142 (15.5)	67 (14.8)
Missing		21 (1.5)	10 (1.1)	11 (2.4)

NOA, non-obstructive azoospermia; TESE, testicular sperm extraction; AFZ, azoospermia factor.

Table II Multivariable logistic regression model for successful sperm retrieval with TESE: stepwise-built logistic model, each row depicting the cumulative contribution of a variable to a model including all variables from previous rows.

Covariate	OR	95% CI	*P*-value	AUC	AUC corrected
Male age	1.06	1.03–1.09	<0.001	0.56	0.56
LH	0.91	0.88–0.93	<0.001	0.63	0.62
LH^2	1.002	1.000–1.003	0.05	0.64	0.63
FSH	0.98	0.96–0.99	0.003	0.65	0.64
Testosterone	1.03	1.01–1.06	0.023	0.65	0.64
AZFc deletion	0.24	0.11–0.52	<0.001	0.66	0.65
Idiopathic NOA	0.44	0.32–0.59	<0.001	0.69	0.67

AUC, area under the curve; CI, confidence interval; OR, odds ratio.

(Fig. 2B). The predicted probability of successful sperm retrieval was compared with the observed sperm retrieval rate in that category. The difference between the mean observed chance and the mean predicted chance was 2.2–9.4%, which indicates a moderate calibration of the prediction model (Table III). Calibration is shown in Fig. 3B. The model showed good calibration below 50% and above 65%. For the predicted sperm retrieval rate between 50 and 65%, a slight overestimation was seen. However, the CIs of the group with a poor predicted chance (<35%) and the group with a moderate predicted change (50–65%) did not overlap, indicating a reliable distinction between these prognostic groups.

Discussion

This study was designed to develop and validate a model to predict successful testicular sperm retrieval in men with NOA. Our prediction model, built after multivariable logistic regression analysis, demonstrated that higher male age, higher values for serum testosterone and lower

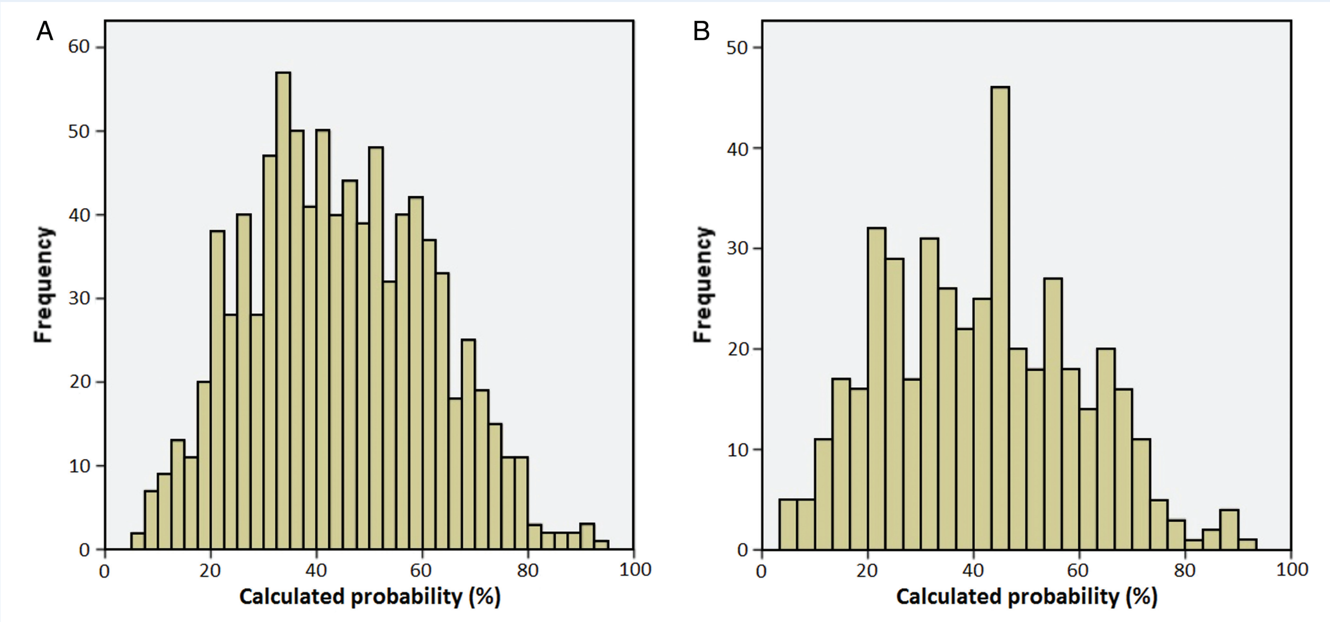

Figure 1 Distribution of the men according to their calculated probabilities of sperm retrieval. (**A**) Development set ($n = 918$) (**B**) Validation set ($n = 453$).

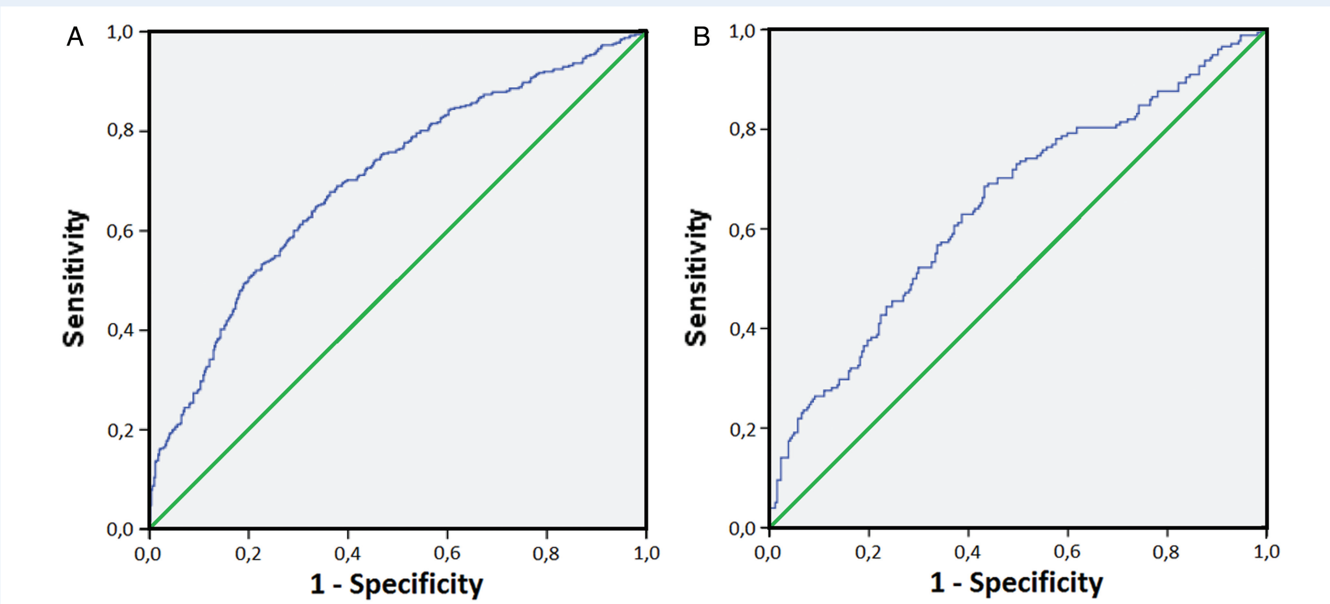

Figure 2 Area under the receiver-operating characteristic curve for prediction of sperm retrieval. (**A**) Development set (**B**) Validation set.

values for serum FSH and LH were predictive for successful sperm retrieval. Diagnosis of idiopathic NOA and the detection of an *AZFc* deletion were predictive for unsuccessful sperm retrieval. The predictive capacity was moderate with an AUC of 0.69 in the development set and 0.65 in the external validation set. The calibration indicated that our model could distinguish men with a poor prognosis from men with a good prognosis, in terms of sperm retrieval.

A number of prediction models have been developed previously to predict successful sperm retrieval with TESE in men with NOA (Samli and Dogan, 2004; Tsujimura *et al.*, 2004; Boitrelle *et al.*, 2011; Ramasamy *et al.*, 2013). The efficacy of these models was moderate. One study reported a sensitivity of 71.0% and a specificity of 71.4% (Tsujimura *et al.*, 2004), while another reported a sensitivity of 68.0% and a specificity of 87.5% (Samli and Dogan, 2004). Two studies performed AUC analysis

Table III Mean predicted probability of successful sperm retrieval versus the mean observed successful sperm retrievals.

Predicted chance	No. of patients in group	Mean predicted chance (%)	No. of successful sperm retrievals	Mean observed chance (%)
Internal validation				
<35%	301	25.4	81	26.9
35–50%	264	42.1	112	42.4
50–65%	228	57.0	134	58.8
>65%	115	72.9	86	74.8
External validation				
<35%	171	23.2	48	28.1
35–50%	134	42.2	53	39.6
50–65%	86	57.1	41	47.7
>65%	51	72.8	36	70.6

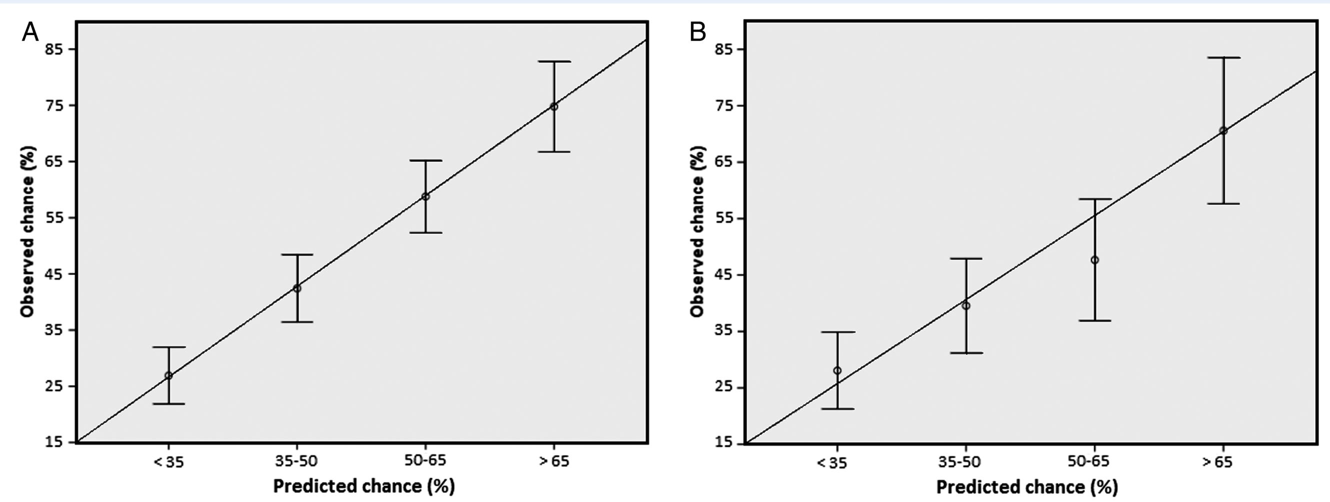

Figure 3 Relationship between calculated and observed sperm retrieval rates. The four groups represent the quintiles of the calculated probabilities. Data on sperm retrieval rate are reported as percentage and 95% confidence interval. (**A**) Development set (**B**) Validation set.

Table IV Five hypothetical men with the calculated probability for obtaining spermatozoa with TESE.

	Man A	Man B	Man C	Man D	Man E
Age (years)	24	30	34	38	42
LH (IU/l)	6	12	8	6	10
FSH (IU/l)	22	26	20	16	20
Testosterone (nmol/l)	20	10	14	18	12
Etiology	Klinefelter syndrome	*AZFc* deletion	Idiopathic NOA	Cryptorchidism	Idiopathic NOA
Calculated probability for obtaining spermatozoa	50%	11%	34%	70%	39%

Probability = 1/[1 + exp(−β)], where β = −1.009 + (male age × 0.058) + (LH − 0.115) + (LH^2 × 0.001) + (FSH − 0.019) + (testosterone × 0.034) + (*AZFc* deletion − 1.480) + (idiopathic NOA − 0.855).

and found an AUC of 0.64 (Ramasamy *et al.*, 2013) and an AUC of 0.66 (Boitrelle *et al.*, 2011). None of the developed models was externally validated. For prediction models in the field of infertility, external validation has a demonstrated lower predictive capacity when evaluated in a different population (Altman and Royston, 2000). Thus, these models need external validation before they can be used in clinical practice.

Four of the independent predictors that we included in our final model were also found previously, i.e. male age, and serum FSH, LH and testosterone levels. Furthermore, previously developed prediction models found that testicular volume, duration of infertility, inhibin B, prolactin, Klinefelter syndrome and history of cryptorchidism were independently correlated with TESE outcome. Most of these variables (not prolactin) were included in our univariable analysis and appeared to have no significant correlation with TESE outcome (duration of infertility) or did not retain their significance after multivariable analysis (testicular volume, inhibin B, Klinefelter syndrome and history of cryptorchidism).

This study has some limitations that need to be addressed. First, there is a potential selection bias, due to the restrictions in government policies on men diagnosed with Klinefelter syndrome. Until September 2010, these men were excluded from undergoing TESE in the Netherlands. The percentage of men with Klinefelter syndrome among men diagnosed with NOA is expected to be higher (11%; Aksglaede and Juul, 2013) than in our study population (4.6% in the development set and 9.5% in the validation set). Secondly, there may be some confounders that affect the chances of finding spermatozoa, such as patient selection, and clinical and laboratory techniques (Vloeberghs *et al.*, 2015). There are no rigorously designed randomized studies that compare the various surgical techniques for effectiveness and safety. A recent meta-analysis, which found that the sperm retrieval rate is higher for micro-TESE compared with conventional TESE and for conventional TESE compared with testicular sperm aspiration (Bernie *et al.*, 2015), also had some serious limitations. The heterogeneity of the population of men diagnosed with NOA was high and consideration of other variables, such as tissue processing techniques, was lacking. Moreover, the factors that we have identified as predictors in sperm retrieval by TESE are likely to be predictive for any of the techniques used to retrieve spermatozoa from the testis. Finally, we did not measure prolactin levels in our study population on a routine basis although it has been mentioned as a possible prognosticator. Of note, previous studies found no differences between prolactin levels in men with a successful sperm retrieval compared with men with unsuccessful sperm retrieval, rendering it unlikely that prolactin is a main predictor for testicular sperm retrieval (Samli and Dogan, 2004; Tsujimura *et al.*, 2004).

A strength of our nationwide cohort is the large number of men that we studied. With a total of 1371 first TESE procedures, it is to our knowledge the largest study on this subject reported thus far. The performance of the external validation of our model further strengthens our study. This study shows that our prediction model can be applied in a general population of men with NOA, without losing its accuracy, regardless of the setting of the hospital.

Once TESE is successful, the second step in achieving a pregnancy is ICSI with the surgically retrieved spermatozoa. Again, predicting the chance of success (in this case live birth) is then important to allow informed decisions to be made. Two studies have attempted to find predictive factors for pregnancy outcome after TESE (Silber *et al.*, 1997; Vernaeve *et al.*, 2004). We recently developed a prediction model where we address the clinical decision to start or continue with ICSI using TESE-derived spermatozoa before treatment or after one or more unsuccessful ICSI treatments (A. M. Meijerink *et al.* submitted for publication). This model is built using the same cohort of patients as this study, with the obvious exception that only those couples that initially underwent a successful sperm recovery with TESE are included in the prediction model for live birth. The model has a moderate discriminative capacity (AUC of 0.62 in the development set; AUC of 0.67 in the validation set) and calibrates well.

Appropriate counselling of men undergoing TESE is an important issue from a psychological viewpoint. Using this model it is possible to counsel men individually about their chance of successfully obtaining spermatozoa with TESE. Considering their individual chance to obtain spermatozoa together with their personal options and preferences will help men and their partners to decide whether to undergo a TESE procedure or not. Men with a poor prognosis might decide to refrain from this treatment option and reconsider artificial insemination of their partner with donor sperm, adoption and/or becoming foster parents to fulfil the child wish. Future studies should clarify patients' preferences and allow clinicians to provide shared decision-making (Den Breejen *et al.*, 2013).

In conclusion, this study demonstrates that in the prediction of successful testicular sperm retrieval a distinction can be made between men with a good prognosis and men with a poor prognosis. The success of sperm retrieval was found to depend on male age, FSH, LH, and testosterone levels, diagnosis of idiopathic NOA and detection of an *AZFc* deletion. After external validation, the model proved to be accurate in predicting chances of success with TESE. We propose that our model should be used for counselling men with NOA.

Authors' roles

All authors proved the quality for authorship by having contributed substantially to this work. All authors developed the original concept of this study collectively. Data collection was performed by M.C., A.M., K.W.D.H. and L.R., statistical analysis by M.C. and 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 by 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 clinican's overview. *Int J Androl* 2010; **33**:e109–e113.

Aksglaede L, Juul A. Testicular function and fertility in men with Klinefelter syndrome: a review. *Eur J Endocrinol* 2013;**168**:R67–R76.

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

Bernie AM, Mata DA, Ramasamy R, Schlegel PN. Comparison of microdissection testicular sperm extraction, conventional testicular sperm extraction, and testicular sperm aspiration for nonobstructive azoospermia: a systematic review and meta-analysis. *Fertil Steril* 2015;**104**:1099–1103.

Bleeker SE, Moll HA, Steyerberg EW, Donders AR, Derksen-Lubsen G, Grobbee DE, Moons KG. External validation is necessary in prediction research: a clinical example. *J Clin Epidemiol* 2003;**56**:826–832.

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.

Bryson CF, Ramasamy R, Sheehan M, Palermo GD, Rosenwaks Z, Schlegel PN. Severe testicular atrophy does not affect the success of microdissection testicular sperm extraction. *J Urol* 2014;**191**:175–178.

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

Chen SC, Hsieh JT, Yu HJ, Chang HC. Appropriate cut-off value for follicle-stimulating hormone in azoospermia to predict spermatogenesis. *Reprod Biol Endocrinol* 2010;**8**:108.

Colpi GM, Piediferro G, Nerva F, Giacchetta D, Colpi EM, Piatti E. Sperm retrieval for intra-cytoplasmic sperm injection in non-obstructive azoospermia. *Minerva Urol Nefrol* 2005;**57**:99–107.

Den Breejen EM, Nelen WL, Schol SF, Kremer JA, Hermens RP. Development of guideline-based indicators for patient-centredness in fertility care: what patients add. *Hum Reprod* 2013;**28**:987–996.

Donoso P, Tournaye H, Devroey P. Which is the best sperm retrieval technique for non-obstructive azoospermia? A systematic review. *Hum Reprod Update* 2007;**13**:539–549.

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.

Hunault CC, Habbema JD, Eijkemans MJ, Collins JA, Evers JL, te Velde ER. Two new prediction rules for spontaneous pregnancy leading to live birth among subfertile couples, based on the synthesis of three previous models. *Hum Reprod* 2004;**19**:2019–2026.

Jarow JP, Espeland MA, Lipshultz LI. Evaluation of the azoospermic patients. *J Urol* 1989;**142**:62.

Jungwirth A, Diemer T, Kopa Z, Dohle GR, Giwercman A, Kopa Z, Krausz C, Tournaye H. Guidelines on male infertility 2015. http://uroweb.org/wp-content/uploads/EAU-Guidelines-Male-Infertility-20151.pdf (March 2014, date last accessed).

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.

Matsumiya K, Namiki M, Takahara S, Kondoh N, Takada S, Kiyohara H, Okuyama A. Clinical study of azoospermia. *Int J Androl* 1994; **17**:140–142.

Ramasamy R, Padilla WO, Osterberg EC, Srivastava A, Reifsnyder JE, Niederberger C, Schlegel PN. A comparison of models for predicting sperm retrieval before microdissection testicular sperm extraction in men with nonobstructive azoospermia. *J Urol* 2013; **189**:638–642.

Samli MM, Dogan I. An artificial neural network for predicting the presence of spermatozoa in the testes of men with nonobstructive azoospermia. *J Urol* 2004;**171**:2354–2357.

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

Silber SJ. Microsurgical TESE and the distribution of spermatogenesis in non-obstructive azoospermia. *Hum Reprod* 2000;**15**:2278–2284.

Silber SJ, Nagy Z, Devroey P, Camus M, Van Steirteghem AC. The effect of female age and ovarian reserve on pregnancy rate in male infertility: treatment of azoospermia with sperm retrieval and intracytoplasmic sperm injection. *Hum Reprod* 1997;**12**:2693–2700.

Toulis KA, Iliadou PK, Venetis CA, Tsametis C, Tarlatzis BC, Papadimas I, Goulis DG. Inhibin B and anti-Mullerian hormone as markers of persistent spermatogenesis in men with non-obstructive azoospermia: a meta-analysis of diagnostic accuracy studies. *Hum Reprod Update* 2010; **16**:713–724.

Tournaye H. Update on surgical sperm recovery—the European view. *Hum Fertil (Camb)* 2010;**13**:242–246.

Tsujimura A, Matsumiya K, Miyagawa Y, Takao T, Fujita K, Koga M, Takeyama M, Fujioka H, Okuyama A. Prediction of successful outcome of microdissection testicular sperm extraction in men with idiopathic nonobstructive azoospermia. *J Urol* 2004;**172**:1944–1947.

Vernaeve V, Krikilion A, Verheyen G, Van Steirteghem A, Devroey P, Tournaye H. Outcome of testicular sperm recovery and ICSI in patients with non-obstructive azoospermia with a history of orchidopexy. *Hum Reprod* 2004;**19**:2307–2312.

Vloeberghs V, Verheyen G, Haentjens P, Goossens A, Polyzos NP, Tournaye H. How successful in 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: Cambridge Press, 2009.

Yang Q, Huang YP, Wang HX, Hu K, Wang YX, Huang YR, Chen B. Follicle-stimulating hormone as a predictor for sperm retrieval rate in patients with nonobstructive azoospermia: a systematic review and meta-analysis. *Asian J Androl* 2015;**17**:281–284.

Yildirim ME, Koc A, Kaygusuz IC, Badem H, Karatas OF, Cimentepe E, Unal D. The association between serum follicle-stimulating hormone levels and the success of microdissection testicular sperm extraction in patients with azoospermia. *Urol J* 2014;**11**:1825–1828.