

RESEARCH

Open Access



# Does the FNA sperm retrieval failure prediction model work well for current NOA individuals undertaking risk screening before the operation? Model validation, high-risk population identification and potential alternative sperm retrieval exploration

Xiaohui Jiang<sup>1,2,3†</sup>, Dingming Li<sup>1,2,3</sup>, Yi Zheng<sup>1,2,3†</sup>, Yinxian Li<sup>4</sup>, Hengzhou Bai<sup>1,2,3</sup>, Guicheng Zhao<sup>1,2,3</sup>, Yi Zhang<sup>5</sup> and Yue Ma<sup>5\*</sup>

## Abstract

**Background** Non-obstructive azoospermia (NOA), the severe type of male infertility. The objective of this study was to evaluate the predictive accuracy of a prediction model of sperm retrieval failure with fine needle aspiration (FNA).

**Methods** This study involved 769 NOA patients (dataset 1) undertaking FNA and 140 NOA patients undertaking mTESE (dataset 2). The previous model was validated and then reconstructed for more potential risk factors and better accuracy in dataset 1. The reconstructed model was evaluated in NOA patients with different new variables. The outcomes of the micro- testicular sperm extraction (mTESE) were compared with the predicted outcomes of FNA to evaluate its potential as an alternative surgical sperm retrieval (SSR) technique.

**Results** 307 (39.92%) males experienced sperm retrieval failure in FNA while 92 (65.7%) males experienced sperm retrieval failure in mTESE. The refined model has 80% overall agreement ( $n=616$ ). The reconstructed model had an AUROC of 0.876 (95% CI: 0.850–0.921). The mTESE has significantly higher success rate (34.29%) than the predicted success rate of FNA (5.71%).

**Conclusions** Previous model shows good consistency. mTESE can be an alternative SSR method for NOA patients with a high predicted risk of sperm retrieval failure with FNA.

**Keywords** Non-obstructive azoospermia (NOA), Fine needle aspiration (FNA), Microtesticular sperm extraction (mTESE), Sperm retrieval risk prediction, Logistic regression

<sup>†</sup>Xiaohui Jiang and Yi Zheng contributed equally to this work.

\*Correspondence:

Yue Ma  
gordonrozen@scu.edu.cn

Full list of author information is available at the end of the article



## Background

Azoospermia is defined as the complete absence of sperm from the ejaculate. This must be confirmed by high-powered microscopic examination of centrifuged seminal fluid on at least two occasions [1]. Azoospermia is divided into obstructive azoospermia (OA) and non-obstructive azoospermia (NOA). As 10–15% of infertile men suffer from azoospermia, over 60% of them are determined to have NOA [2]. NOA is caused by testicular spermatogenesis dysfunction, sex chromosome abnormalities, Y chromosome micro-deletions, etc [3–5]. NOA, as the severe type of male infertility, has become a serious challenge to male reproductive health.

NOA patients can obtain sperm through surgical sperm retrieval (SSR) and have their own biological offspring via intracytoplasmic sperm injection (ICSI) technology. SSR methods include fine needle aspiration (FNA), testicular sperm extraction (TESE) and micro-testicular sperm extraction (mTESE), the latest SSR method [6]. Compared to other SSR methods, FNA is less invasive to obtain a part of mapping for histological diagnosis while obtaining enough sperm to meet the needs of ICSI and thus has been partly used in reproductive medicine centers. In general, NOA patients who aim to have their own biological offspring usually explore SSR options first and only consider donor sperm-assisted pregnancy in the case of failure of mTESE or other SSR options [7].

The sperm retrieval rate (SRR) of FNA is approximately 11–44% in NOA patients; therefore, NOA patients should be well informed of the risk of sperm retrieval failure associated with FNA before their in vitro fertilization treatments [8]. On the other hand, the risk of FNA is minimal compared to other forms of SSR and this is why it is better to consider before going for a more invasive procedure [9]. Importantly, the ability to predict the outcome of FNA would be of great value in both counseling patients [10] and clinical applications.

Some studies have tried to predict the outcome of sperm retrieval. However, the models developed thus far cannot accurately predict sperm retrieval, and few of them have been sufficiently validated. Previously, we have explored the potential influencing factors of sperm retrieval failure with FNA in NOA patients and tried to build a new prediction model [11]. In that study, 327 males with NOA, about 54.8% of patients tested, experienced sperm retrieval failure with FNA. The follicle-stimulating hormone (FSH) level, age and testicular volume were included to establish the prediction model for sperm retrieval failure risk. The prediction accuracy of the model was approximately 85% in external validation, while the external prediction model of previous research report was only 67% in the Netherlands [12].

Since then, we have applied our predictive model to screen out NOA patients who are highly likely to

experience sperm retrieval failure with FNA and to recommend alternative sperm retrieval methods such as mTESE. After such screening, our population of NOA patients undertaking FNA has been changed from an unselected FNA population to a specific population with a lower predicted risk of sperm retrieval failure. Therefore, it is necessary to advance further research to evaluate the predictive accuracy of our model for further clinical applications. Moreover, the consistency of the association between FNA outcome and related factors in current NOA patients undertaking FNA also needs to be explored in order to provide clues about the potential mechanism of FNA failure.

The aim of this study was to validate and improve the predictive accuracy of the existing sperm retrieval failure risk prediction model described in our previous study and to determine whether mTESE would be a potential solution for NOA patients with a high risk of sperm retrieval failure. These findings would provide evidence for making SSR decisions for NOA patients with different expected risks of FNA failure.

## Materials and methods

### Data collection

#### *Patient population*

We collected data of 909 NOA patients who visited the Andrology Center of West China Second Hospital, Sichuan University from October 2018 to November 2021, including age, infertility duration, testicular volume, reproductive hormone (including follicle-stimulating hormone, luteinizing hormone, estrogen, and testosterone) levels, chromosome karyotype, epididymal status, and Johnsen score. Patients who met all the following criteria were included: [1] no hypogonadotropic hypogonadism; [2] no history of cancer and related treatment; [3] never used drugs that alter sex hormone levels (e.g., exogenous testosterone, gonadotropins, aromatase inhibitors, selective estrogen receptor modulators) within 3 months; [4] no history of testicular trauma; [5] no obstructive factors.

According to the semen analysis results (at least twice), characteristics of the medical history, sex hormone levels, and ultrasound results of the reproductive system, NOA was defined after excluding known obstructive factors. Testicular volume was measured with a Prader orchimeter due to limited conditions. Reproductive hormones were measured by chemiluminescence analysis of venous blood collected between 8:00 and 11:00 a.m. after an overnight fast. The status of the epididymis was determined by ultrasound of the reproductive system to determine whether there were structural abnormalities. The Johnsen score was determined by analysis of testicular tissue by an experienced pathologist.

The protocol for this study was approved by the Ethics Committee of West China Second University Hospital of Sichuan University. The project number was 2,018,028. All couples signed informed consent for treatment and follow-up before participating in this study.

#### **FNA procedure**

We performed blind testicular puncture through the skin with FNA. After routine disinfection and draping, the patient was positioned in a supine position, and the procedure was performed under local anesthesia. The testicle with the relatively larger testicular volume or the right testicle in cases of equal testicular volume was selected as the puncture site to avoid a left invisible varicocele. Then, a spermatic nerve block with 5 ml of 2% lidocaine hydrochloride was performed on the surgical side, infusing anesthesia into the testicular tunica albuginea for ~3 min. The operator fixed the testicle of the patient with the left hand and held a 20-ml side-hole needle with the right hand to puncture the testicular albuginea after anesthesia was effective. When the needle reached the appropriate depth to sustain negative pressure for aspiration of testicular tissue and resistance was felt, the needle was slowly withdrawn. The testicular tissue was aspirated through the scrotal skin puncture point, and the needle was completely recovered. Then, haemostasis was carefully evaluated to confirm no visible bleeding at the puncture site. The seminiferous tubules were separated with a needle in the culture fluid drop (Tyrode's fluid, CAF 2.5 mM PTX 7.5 mM pH 7.4) on a slide to observe their appearance, thickness and fullness under a dissecting microscope ( $\times 25$ ). Punctured seminiferous tubules were removed, and the cell suspension was covered with a coverslip. The morphology and activity of mature sperm were observed under an inverted microscope ( $\times 400$ ). If there were mature spermatozoa, ~100 sperm stained with eosin were counted to calculate the sperm survival rate.

#### **mTESE procedure**

All mTESE procedures were performed by the same team of 4 andrologists. The procedure based on the described technique previously. Briefly, one mid-pole sagittal incision in the tunica albuginea was made to allow visualization of the testicular parenchyma without affecting the testicular blood supply under the operating microscope magnified 10 times. Direct examination of the testicular parenchyma was then performed with a surgical microscope at magnification of 20 times. Small samples were excised containing tubules that were particularly large and opaque. Then laboratory staff searched for sperm under a microscope with a magnification of 200 times and tried to freeze them. Surgical termination was defined as freezing enough mature sperm for ICSI.

Testicular histopathological specimens with a diameter of 0.5 mm were obtained from all patients during the operation randomly and analyzed by a professional urogenital pathologist.

#### **Study design**

This study was carried out in three phases. In the first phase, the characteristics of the current NOA patients who undertaken FNA were presented to reveal the population change after screening using the predicted risk of sperm retrieval failure. At this stage, the dataset 1 was collected to refine the risk prediction model proposed in our previous work, which predicts the risk of sperm retrieval failure in the NOA individuals.

The NOA individuals in dataset 1 had taken the risk screening before inclusion, so most of them did not have high risks of sperm retrieval failure. The related analysis is carried out in phase two. In the second phase, the accuracy and identified risk factors for the risk prediction model built in our earlier work were validated in the current NOA population that undertaken FNA. Then, a new prediction model was built for the current NOA populations. The performance of the prediction model was then compared to the former model built for the former NOA populations without screening according to the predicted risk of sperm retrieval failure. Moreover, to clarify the performance of the optimal model in different populations, the prediction accuracy was evaluated in the current NOA patients with specific clinical features, such as different Johnsen scores, chromosomal statuses and epididymis status.

In the third phase, a group of NOA patients screened out due to a high predicted risk of FNA sperm retrieval failure were recommended to undertake mTESE, and this part of patients constitutes dataset 2. Their mTESE outcomes were collected and compared to the predicted outcomes to validate whether mTESE would be a proper alternative treatment for NOA patients who are highly likely to experience sperm retrieval failure with FNA.

#### **Statistical analysis**

##### ***Clinical characteristics and potential predictor identification***

The demographic and clinical characteristics of the current NOA patients with FNA were described. The potential predictors between the NOA patients with successful and failed sperm retrieval with FNA were tested.

The variables with normal distributions were tested using t tests, and other continuous variables were tested using nonparametric tests. Categorical variables were tested using chi-square tests. The test level was set to 0.1 for a higher identification rate.

**Model validation, optimization, and comparison**

To evaluate the prediction accuracy of the existing model (referred to as model 1 hereinafter) in the current NOA patients with lower predicted risks of sperm retrieval failure, the overall agreement between predicted and observed outcomes was calculated by using dataset 1. Moreover, to validate whether the risk factors found in model 1 were consistent, one other logistic risk prediction model with the same variables as model 1 was built (referred to as model 2 hereinafter) using dataset 1. Then, its receiver operating characteristic (ROC) curve and regression coefficients were compared with those of model 1.

To improve the predictive accuracy, a new logistic risk prediction model (model 3) was built based on dataset 1 in a stepwise manner, and the criteria for inclusion and exclusion were 0.10 and 0.05, respectively. The ROC curve of model 3 was built to calculate the AUC and the optimal cutoff value, with which the overall agreement, specificity and sensitivity were calculated and then compared to those of model 1. The variables included were compared for further analysis.

The process of model establishment in this study is shown in Fig. 1.

**Subgroup analysis**

According to variable comparisons between model 2 and model 3 and the findings of other studies, subgroup analyses were carried out between the NOA patients with

different Johnsen scores, chromosomal abnormalities and other risk factors identified in dataset 1.

**Successful rate comparison between sperm retrieval using FNA and mTESE**

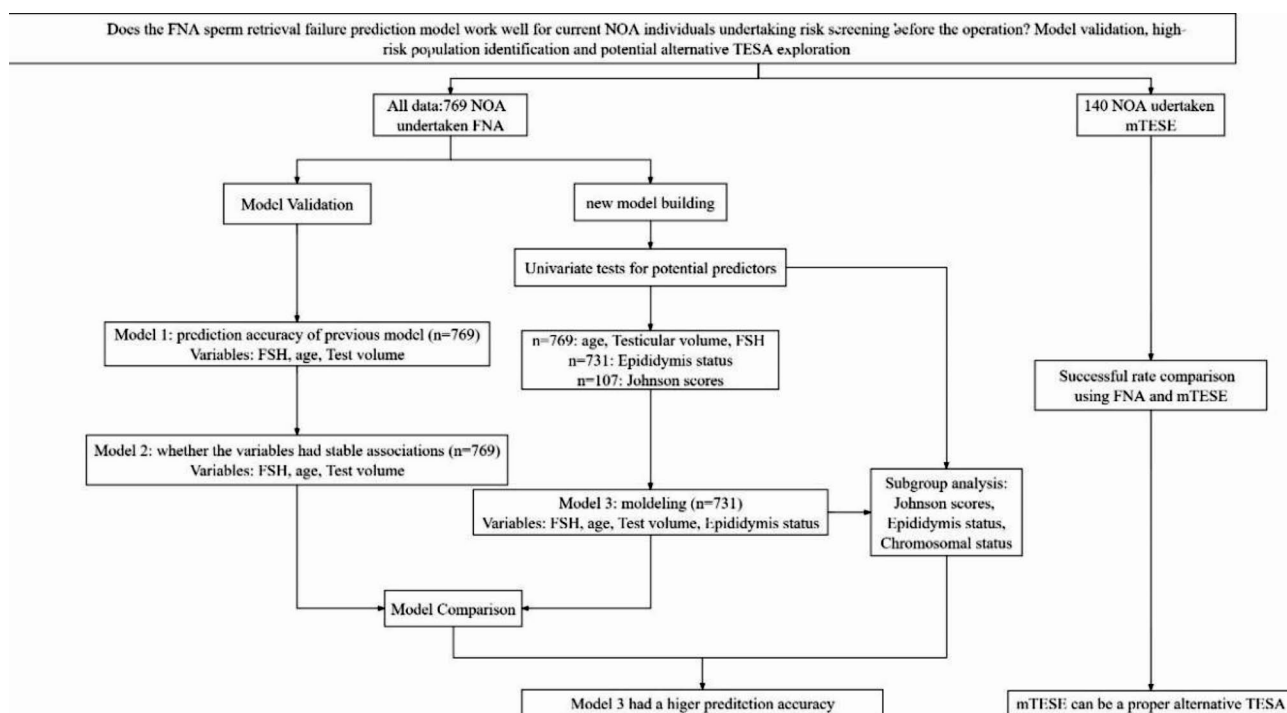
According to the predictors of the NOA patients who had undertaken mTESE, the outcomes of FNA were predicted and compared with the observed outcomes of mTESE. The success rates of sperm retrieval were calculated and tested using the chi-square test.

**Results**

In the dataset 1, the average age of the 769 patients was 30 years old. There were 307(39.92%) males undertaking FNA and experiencing sperm retrieval failure. In the dataset 2, the average age of the 140 patients was 30 years old. There were 92 (65.7%) males undertaking mTESE and experiencing sperm retrieval failure. The baseline characteristics are shown in Table 1.

**Potential predictors of sperm retrieval failure with FNA**

The univariate tests showed that the differences in age, testicular volume, FSH level, Johnsen score, and epididymis status were significant between the NOA patients with successful and failed sperm retrieval (Table 2).



**Fig. 1** Flow chart of model establishment in this study

**Table 1** Baseline characteristics of NOA patients in the two datasets

	Dataset 1	Dataset 2
	Mean ± SD or median (IQR)	Mean ± SD or median (IQR)
Age (year)	29.0(27.0,32.0)	30.0(27.0,32.0)
FSH (IU/l)	6.7(4.0,14.6)	26.90(20.18,37.50)
Infertility time (year)	2.0(1.0,3.0) in 714	
Testicular volume (ml)	12.0(10.0,16.0)	5.3(2.6,7.1)
Estrogen (pg/ml)	30.0(23.9,37.6) in 742	28.25(20.32,34.52)
Testosterone (ng/ml)	3.5(2.6,4.8) in 767	2.91(1.78,3.85)
LH (IU/l)	4.1(2.8,6.0) in 753	9.60(6.18,17.50)
Johnsen score	2.0(2.0,4.0) in 107	1.00(1.00,1.75)
Primary infertility (% , n)	(91.86%, 666)	
Epididymis abnormality (% , n)	(23.12%, 169)	
Chromosomal abnormality (% , n)	(5.79%, 7) in 121	
Sperm retrieval failure (% , n)	(39.92%, 307)	(65.7%, 92)

dataset 1: n1 = 769 dataset 2: n2 = 140

**Table 2** Univariate tests for potential predictors

	Wilcoxon W	χ-square	P
Age (years)	94,598		< 0.001
Testicular volume (ml)	48,334		< 0.001
FSH (IU/l)	123,955		< 0.001
Johnsen score	0		< 0.001
Epididymis status		20.074	< 0.001

**Model validation, modeling, and comparison**

**Model validation**

The previously built model 1 was used to calculate the predicted risk of sperm retrieval failure for all 769 NOA patients, and the cutoff value (0.6461) was used to judge the result of predicted sperm retrieval. The overall agreement of the predicted and observed outcomes was 80% (n=616). The predicted risk of sperm retrieval failure of model 1 was calculated as follows:

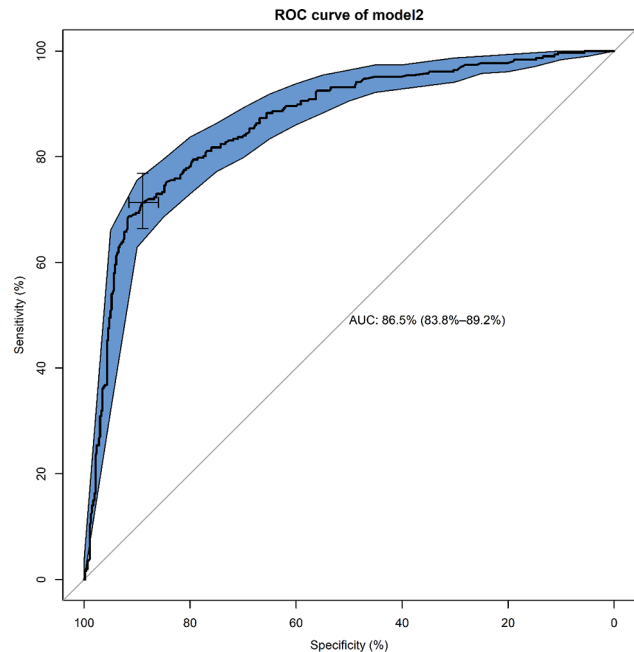
$$\begin{aligned}
 B = & 0.003 + (0.124 \times FSH) \\
 & + (-1.448 \times \text{Age } 35 \text{ years or older}) \\
 & + (-0.651 \times \text{Test volume } 12 - 14 \text{ ml}) \\
 & + (-1.448 \times \text{Test volume } 15 \text{ ml or greater})
 \end{aligned}$$

Logistic regression modeling (model 2): To test whether the variables in model 1 in the 769 NOA patients had stable associations with FNA sperm retrieval outcomes, model 2 was built, including age, testicular volume, and FSH level, which were used in model 2 as shown below:

**Table 3** Risk prediction model 2

Variable	Coefficient (95% CI)	P	OR (95% CI)
Constants	-1.212(-1.690,-0.747)	< 0.001	0.298 (0.184,0.474)
Testicular volume: 12–14 ml	-0.374(-0.816,-0.068)	0.097	0.688 (0.442,1.070)
Testicular volume: >=15 ml	-1.495(-1.977,-1.020)	< 0.001	0.224 (0.138,0.360)
FSH (IU/l)	0.164 (0.133,0.196)	< 0.001	1.178 (1.142,1.217)
Age (years)	-1.592(-2.280,-0.955)	< 0.001	0.204 (0.102,0.385)

N=769



**Fig. 2** ROC curve analysis of prediction model 2

$$\begin{aligned}
 B = & -1.212 + (0.164 \times FSH) \\
 & + (-1.592 \times \text{Age } 35 \text{ years or older}) \\
 & + (-0.374 \times \text{Test volume } 12 - 14 \text{ ml}) \\
 & + (-1.495 \times \text{Test volume } 15 \text{ ml or greater})
 \end{aligned}$$

The model suggested that a low FSH level, a large testicular volume, and an age of 35 years or older were associated with a low risk of FNA failure (Table 3).

The corresponding ROC (Fig. 2) curve had an AUC of 0.865 (95% CI: 0.838–0.892). Youden’s index showed that the best cutoff value for predicting the risk of sperm retrieval failure was 0.3746, with a sensitivity of 0.79478 and a specificity of 0.79437.

Logistic regression modeling (model 3): To achieve possible improvement in the prediction accuracy, a new logistic regression risk prediction model (model 3) was built as follows:

$$\begin{aligned}
 B = & -1.140 + (0.167 \times FSH) \\
 & + (-1.438 \times \text{Age 35 years or older}) \\
 & + (-0.391 \times \text{Test volume 12 - 14 ml}) \\
 & + (-1.451 \times \text{Test volume 15 ml or greater}) \\
 & + (-0.870 \times \text{epididymis status})
 \end{aligned}$$

The included variables were age, testicular volume, FSH level, and epididymis status. A low FSH level, a large testicular volume, a normal epididymis, and an age of 35 years or older were associated with a low risk of sperm retrieval failure (Table 4).

The corresponding ROC (Fig. 3) curve had an AUC of 0.876 (95% CI: 0.850–0.921). Youden’s index showed that the best cutoff value for predicting the risk of sperm retrieval failure was 0.391, with a sensitivity of 0.793 and a specificity of 0.825.

**Model comparison**

In the three models, the coefficients had the same directions except for the intercept (Table 5). There were no significant differences in the coefficients and AUCs between model 2 and model 3 based on dataset 1. In addition, the absolute value of the coefficient of testicular volume, 12–14 ml, in model 1 was obviously smaller than those in model 2 and model 3.

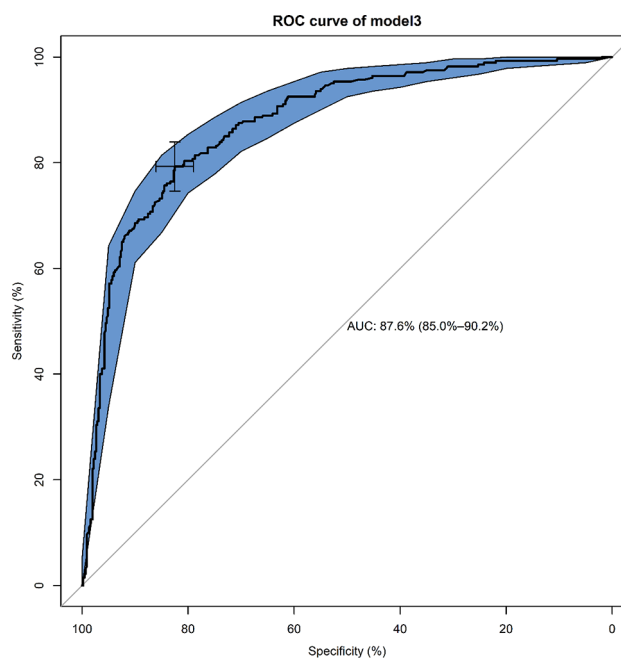
**Subgroup analysis**

According to clinical experience and related research, the Johnsen score was used to define two groups ( $\leq 2$  and  $> 2$ ) to explore the influence on the risk of sperm retrieval. The results suggested that the consistency rate between the predicted outcomes and the observed outcomes was 43.33% in group 1 ( $\leq 2$ ), which was significantly smaller ( $p < 0.05$  by t test) than that of group 2 ( $> 2$ ) (95.16%). For chromosomal status, the consistency rate was 82.46% in group 1 (normality), which was not significantly different ( $p > 0.05$  by t test) from that of group 2 (abnormality) (85.71%). Model 3 suggested that epididymis status was related to the risk of sperm retrieval failure, which was not found in an earlier study (model 1). Thus, it was necessary to conduct a subgroup analysis of epididymis status to explore the risk of sperm retrieval failure in males with epididymis normality or abnormality. The results suggested that the consistency

**Table 4** Risk prediction model 3

Variable	Coefficient (95% CI)	P	OR (95% CI)
Constants	-1.140 (-1.670, -0.621)	< 0.001	0.320 (0.188, 0.537)
Testicular volume: 12–14 ml	-0.391 (-0.867, 0.084)	0.107	0.676 (0.420, 1.087)
Testicular volume: $\geq 15$ ml	-1.451 (-1.963, -0.947)	< 0.001	0.232 (0.139, 0.384)
FSH	0.167 (0.135, 0.201)	< 0.001	1.182 (1.144, 1.223)
Epididymis status	-0.870 (-1.388, -0.375)	< 0.001	0.419 (0.250, 0.687)
Age	-1.438 (-2.144, -0.787)	< 0.001	0.237 (0.117, 0.455)

N=769



**Fig. 3** ROC curve analysis of prediction model 3

rate between the predicted outcomes and the observed outcomes was 78.11% in group 1 (normality), which was significantly smaller ( $p < 0.05$  by t test) than that of group 2 (abnormality) (88.17%).

**Successful rate comparison of sperm retrieval with FNA and mTESE**

As the variable epididymis status was not examined before mTESE, model 3 could not be used to evaluate

**Table 5** Comparison of model coefficients

Variable	Model 1	Model 2	Standardized coefficient	Model 3	Standardized coefficient
	coefficient	coefficient		coefficient	
Intercept	0.003	-1.212	0	-1.140	0
Testicular volume: 12–14 ml	-0.651	-0.374	-0.341	-0.391	-0.361
Testicular volume: $\geq 15$ ml	-1.382	-1.495	-1.501	-1.451	-1.475
FSH	0.124	0.164	2.769	0.167	2.835
Epididymis status	\	\	\	-0.870	-0.754
Age	-1.448	-1.592	-1.125	-1.438	-1.034

the effects of sperm retrieval technologies. Thus, model 1 and model 2 were employed to obtain the expected outcome using predicted risks.

According to model 1, if the 140 patients had undertaken FNA instead of mTESE, only 8 were expected to have successful sperm retrieval. Furthermore, 48 of these patients who undertaken mTESE experienced successful sperm retrieval. The results suggested that the success rate of FNA was 5.71%, which was significantly lower ( $p < 0.05$  by chi-square test) than that of mTESE (34.29%).

## Discussion

In our recent study, we proposed a predictive model that we have since used to assess the sperm retrieval failure risk in our current NOA population before undertaking FNA; based on this evaluation, individuals who are highly likely to experience FNA failure are recommended to undertake an alternative SSR process. This study validated the risk prediction model built earlier in current NOA population that undertaken FNA, evaluated the consistency of correlations between identified risk factors and sperm retrieval failure and explored whether mTESE is an appropriate alternative SSR for NOA individuals with a high risk of sperm retrieval failure.

To date, FNA is still one kind of common SSR technique. FNA was first described to assess male fertility in 1965 [13]. The fine needle specifications and puncture points used in this technology have been gradually standardized in the past 50 years [14, 15]. Globally, FNA, TESE and mTESE have become important surgical methods for sperm retrieval. Compared with TESE, FNA has a similar SRR with less testicular injury and fewer complications [16]. FNA plays an irreplaceable role as a surgical method because of its diagnostic value and lower cost in the era of microsurgery [17].

As before, for NOA individuals undertaking FNA, we built a sperm retrieval failure risk prediction model that outperformed existing models. The model included three independent predictors: serum FSH level, testicular volume and age. Subsequently, NOA patients with high risk factors or a high predicted risk of sperm retrieval failure were recommended to undertake alternative SSR processes for a higher sperm retrieval rate. This led to a change in the population of NOA individuals who undertaken FNA, e.g., some of the high-risk NOA patients undertaken other processes, while some of these individuals still undertaken FNA due to the fewer side effects and smaller economic burden associated with this technique or other personal factors. Whether the risk prediction model still works in such a changing population of NOA individuals undertaking FNA is unclear.

In the prospective cohort of 769 NOA individuals undertaking FNA, 38.3% of the NOA individuals experienced sperm retrieval failure with FNA, which was

54.8% in the NOA population without risk screening in the previous study. This considerable difference in the sperm retrieval failure rate with FNA suggested that risk screening based on predicted risk substantially changed the population of NOA patients who undertaken FNA. With the validation in the prospective cohort, the earlier built model maintained a high prediction accuracy of 80.0%, even with the considerable decrease in the high-risk individuals.

Along with the validation of the earlier prediction model, the consistency of the association between the predictors and sperm retrieval failures was evaluated. The results suggested that the predictors maintained similar correlations, with both intensity and direction, to sperm retrieval failure in the changing population of NOA individuals undertaking FNA. Such robustness of the associations suggests that specific mechanistic pathways may exist in those predictors.

In early studies, the serum FSH level was confirmed to be inversely proportional to the number of spermatogonia and primary spermatocytes in the testis [18]. A lower FSH level has been shown to reflect sufficient feedback from the testicular germ cells and Sertoli cells. Sertoli cells secrete testosterone-dependent paracrine stimuli for germ cells, and FSH targets Sertoli cells through FSH receptors (FSHRs) during spermatogenesis [19]. FSH is currently used as a possible predictor of the outcome of NOA patients.

Many studies have reached opposite conclusions about the predictive value of FSH, especially in the prediction of the outcomes of TESE and mTESE [20–22]. However, another study on FNA obtained more satisfactory results [23]. It can be inferred that different prediction models should be built for corresponding sperm retrieval methods.

In addition, other factors associated with sperm retrieval failure were explored in the changing NOA population undertaking FNA. A novel prediction model including epididymis status exhibited a slight improvement in prediction accuracy. The normal epididymis status is correlated with a lower risk of sperm retrieval failure. The male epididymis is formed by the confluence of many interconnected efferent tubules, which store epididymal fluid and sperm. The epididymis may reflect the state of spermatogenesis inside the testis from the side to a certain extent. Pezzella et al. [24] found that in NOA patients, a smaller epididymal head diameter is associated with a smaller testicular volume and higher serum FSH level. This may suggest that if the anatomical structure of the epididymis is abnormal or affected by infection factors, even if spermatogenesis is normal in the testis, it may affect the normal passage of sperm. When an epididymal infection spreads backwards into the testis, it may affect spermatogenesis in the testis, and in

severe cases, it may even lead to the risk of NOA. Thus, there could be a potential relationship between the epididymis and SRR and physical examination from the external reproductive system may reveal enlargement of the epididymis or nodular changes in the tail of the epididymis or even soft testis. However, due to the incompleteness of the previous data, the epididymis factor has not been explored in previously constructed models. To this end, our new study, which combines medical history collection, physical examination, and imaging evaluation, provides a more detailed description, and defines the epididymal status. Men whose physical examination revealed a deformed epididymis and nodular changes confirmed by ultrasound were defined as patients with an abnormal epididymal status. The status of the epididymis was included as a factor that may be associated with the SRR in the analysis. This may suggest that an abnormal epididymal status should be clearly assessed before FNA sperm retrieval.

The testicular volume was divided into different groups with 15 ml as the boundary. The volume associated with “small testes” is 12–15 ml. However, our study involved East Asian individuals, and there are ethnic differences compared with other related studies worldwide. Previous studies suggest that the hypotrophic testicular volume in East Asian individuals is slightly lower than that in white European and American individuals [25]. This indicates that the cutoff value of the risk factor for “testicular volume” needs to be adjusted for East Asian individuals, and 12 ml was found to be a more appropriate cutoff value, at least for predicting sperm retrieval failure with FNA.

Moreover, a small number of subjects, 92 and 121, had Johnsen scores and chromosome abnormality test results, respectively. Subgroup analysis for these individuals suggested that the prediction model worked well only for NOA individuals with Johnsen scores greater than 2. Additionally, such differences were found in NOA individuals with different epididymis statuses. The prediction model did not achieve satisfactory accuracy when subjects had an abnormal epididymis status. However, no significant difference for NOA individuals with abnormal chromosomes was found in the subgroup analysis. These subgroup analyses suggested that this prediction model should not be employed to assess the risk of sperm retrieval failure with FNA in NOA patients with Johnsen scores less than 2 or an abnormal epididymis status.

For NOA patients who are highly likely to have sperm retrieval failure, alternative SSR techniques with higher expected success rates are recommended. A total of 140 of these patients decided to undertake mTESE. Based on the risk prediction model of sperm retrieval failure, it was expected that only 8 of these patients would have had successful sperm retrieval with FNA. However, 48 of the 140 experienced successful sperm retrieval with mTESE.

This confirmed that mTESE can be a proper alternative SSR technique with a higher success rate for NOA patients with a high predicted risk of sperm retrieval failure with FNA. Such techniques are capable of retrieving sperm in NOA patients for whom it is difficult to retrieve sperm with FNA. The advantage of mTESE compared to single- or multipoint FNA is that it can distinguish thicker and full seminiferous tubules, in which developed sperm can be obtained, by using a high-power microscope. Furthermore, at the technical level, mTESE may be more suitable for NOA patients with small testes, especially those with an average unilateral testis volume less than 8 ml, including the vast majority of Klinefelter syndrome patients, which are the most common NOA patients with chromosomal abnormalities. Similarly, for NOA patients with large testes and normal gonadal development, as the seminiferous tubules in the testes may be generally thick, FNA is suitable as the preferred method of sperm retrieval due to the advantages of lower cost, less injury, and easier operation. Research has also shown that the rational selection of the acquisition technology through the sequential method is an important strategy to improve the SRR [26, 27]. However, due to the 140 NOA patients lacking the sperm retrieval outcome of FNA, the prediction accuracy of the FNA model was not practically validated in this dataset.

As validated in this study, the prediction model we previously built maintains high prediction accuracy in the current NOA population that undertaken FNA, even with substantial change. Additionally, the three predictors showed robust associations with sperm retrieval failures. More detailed research should be carried out on the potential pathway of the mechanism of sperm retrieval failure involving these factors. In addition, according to the comparison of expected outcomes after FNA and observed outcomes after mTESE, NOA patients should first undertake sperm retrieval failure evaluation with the assistance of the prediction model. Then, the inquiry about the proper choice of SSR can be made by clinicians based on the evaluation.

The purpose of this study was to prospectively verify the external validity of our previously established model and determine whether the model is still effective if the population of NOA patients receiving FNA changed after model screening. Finally, we came to a positive conclusion. Therefore, we recommend that this model should be validated in larger-scale and multicenter studies to promote its clinical application as soon as possible. Therefore, we developed the small procedures of FNA sperm retrieval prediction model (<http://score.yunyimd.com/>) to proceed larger-scale and multicenter studies for improving and optimizing the clinical application value of the model.



## Abbreviations

NOA	Non-obstructive azoospermia
FNA	Fine needle aspiration
mTESE	Microtesticular sperm extraction
SSR	Surgical sperm retrieval
OA	Obstructive azoospermia
ICSI	Intracytoplasmic sperm injection
SRR	Sperm retrieval rate
FSH	Follicle-stimulating hormone
LH	Luteinizing hormone
AMH	Anti-Müllerian hormone
ROC	Receiver operating characteristic

## Acknowledgements

None.

## Author contributions

All authors qualify for authorship by contributing substantially to this article. Y. Ma and X. Jiang developed the original concept of this study collectively. Data collection was performed by D. Li, Y. Zheng, Y. Li, H. Bai, and G. Zhao, statistical analysis by Y. Ma and Y. Zhang. All authors have contributed to critical discussion and reviewed the final version of the article and approve it for publication.

## Funding

Science and Technology project of the Health Planning Committee of Sichuan (20PJ085) and Sichuan Science & Technology Program (2021YFS0181). The authors are grateful for the support of the National Natural Science Foundation of China (Grant Nos. 82373689), the Sichuan Science & Technology Program (Grant Nos. 2021YFS0181) and the Chengdu Science & Technology Program (Grant Nos. 2024-YF05-00644-SN).

## Data availability

No datasets were generated or analysed during the current study.

## Declarations

### Ethics approval and consent to participate

The protocol for this multicenter study was approved by the Ethics Committee of West China Second University Hospital of Sichuan University. The project number was 2018028. All couples signed informed consent for treatment and follow-up before participating in this study. All procedures performed in this study were in accordance with the ethical standards of the institutional and/or national research committee(s) and with the Helsinki Declaration (as revised in 2013).

### Consent for publication

Not applicable.

### Competing interests

The authors declare no competing interests.

### Author details

<sup>1</sup>Human Sperm Bank, West China Second Hospital of Sichuan University, Chengdu, Sichuan 610041, China

<sup>2</sup>Key Laboratory of Birth Defects and Related Diseases of Women and Children (Sichuan University), Ministry of Education, Chengdu, Sichuan 610041, China

<sup>3</sup>NHC Key Laboratory of Chronobiology (Sichuan University), Chengdu, Sichuan, China

<sup>4</sup>Chongqing Changshou Women and Children's Hospital, Changshou, Chongqing, China

<sup>5</sup>West China School of Public Health and West China Fourth Hospital, Sichuan University, Chengdu, Sichuan 610041, China

Received: 17 November 2023 / Accepted: 10 December 2024

Published online: 30 December 2024

## References

1. WHO. WHO laboratory manual for the examination and processing of human semen 2010.
2. Hwang K, Smith JF, Coward RM, Penzias A. Evaluation of the azoospermic male: a committee opinion. *Fertil Steril*. 2018;109(5):777–82.
3. Gudeman SR, Townsend B, Fischer K, Walters RC, Crain D. Etiology of azoospermia in a military population. *J Urol*. 2015;193(4):1318–21.
4. Ezech UI. Beyond the clinical classification of azoospermia: opinion. *Hum Reprod*. 2000;15(11):2356–9.
5. Raman JD, Schlegel PN. Testicular sperm extraction with intracytoplasmic sperm injection is successful for the treatment of nonobstructive azoospermia associated with cryptorchidism. *J Urol*. 2003;170(4 Pt 1):1287–90.
6. Carpi A, Sabanegh E, Mechanick J. Controversies in the management of nonobstructive azoospermia. *Fertil Steril*. 2009;91(4):963–70.
7. Gao X, Sun S, Xie L, Lu S. Effects of donor sperm on perinatal and neonatal outcomes resulting from in vitro fertilization-intracytoplasmic sperm injection and embryo transfer cycles: a retrospective cohort study. *Annals Translational Med*. 2022;10(15).
8. 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(5).
9. Eliveld J, van Wely M, Meißner A, Repping S, van der Veen F, van Pelt AMM. The risk of TESE-induced hypogonadism: a systematic review and meta-analysis. *Hum Reprod Update*. 2018;24(4):442–54.
10. Abdel Raheem A, Garaffa G, Rushwan N, De Luca F, Zacharakis E, Abdel Raheem T, et al. Testicular histopathology as a predictor of a positive sperm retrieval in men with non-obstructive azoospermia. *BJU Int*. 2013;111(3):492–9.
11. Ma Y, Li F, Wang L, Zhao W, Li D, Xian Y et al. A risk prediction model of sperm retrieval failure with fine needle aspiration in males with non-obstructive azoospermia. 2019;34(2):200–8.
12. Cissen M, Meijerink AM, D'Hauwers KW, Meissner A, van der Weide N, Mochtar MH, et al. Prediction model for obtaining spermatozoa with testicular sperm extraction in men with non-obstructive azoospermia. *Hum Reprod*. 2016;31(9):1934–41.
13. Obrant KO, Persson PS. [Cytological study of the testis by aspiration biopsy in the evaluation of fertility]. *Urol Int*. 1965;20(3):176–89.
14. Dajani YF, Kilani Z. Role of testicular fine needle aspiration in the diagnosis of azoospermia. *Int J Androl*. 1998;21(5):295–300.
15. Rosenlund B, Kvist U, Plöen L, Rozell BL, Sjöblom P, Hillensjö T. A comparison between open and percutaneous needle biopsies in men with azoospermia. *Hum Reprod*. 1998;13(5):1266–71.
16. Janosek-Albright KJC, Schlegel PN, Dabaja AA. Testis sperm extraction. *Asian J Urol*. 2015;2(2):79–84.
17. Meng MV, Cha I, Ljung BM, Turek PJ. Relationship between classic histological pattern and sperm findings on fine needle aspiration map in infertile men. *Hum Reprod*. 2000;15(9):1973–7.
18. de Kretser DM, Burger HG, Hudson B. The relationship between germinal cells and serum FSH levels in males with infertility. *J Clin Endocrinol Metab*. 1974;38(5):787–93.
19. Shiraishi K, Matsuyama H. Gonadotropin actions on spermatogenesis and hormonal therapies for spermatogenic disorders [Review]. *Endocr J*. 2017;64(2):123–31.
20. Ramasamy R, Lin K, Gosden LV, Rosenwaks Z, Palermo GD, Schlegel PN. High serum FSH levels in men with nonobstructive azoospermia does not affect success of microdissection testicular sperm extraction. *Fertil Steril*. 2009;92(2):590–3.
21. Caroppo E, Colpi EM, D'Amato G, Gazzano G, Colpi GM. Prediction model for testis histology in men with non-obstructive azoospermia: evidence for a limited predictive role of serum follicle-stimulating hormone. *J Assist Reprod Genet*. 2019;36(12):2575–82.
22. Bromage SJ, Falconer DA, Lieberman BA, Sangar V, Payne SR. Sperm retrieval rates in subgroups of primary azoospermic males. *Eur Urol*. 2007;51(2):534–9; discussion 9–40.
23. Goulis DG, Tsamietis C, Iliadou PK, Polychronou P, Kantartzis PD, Tarlatzis BC, et al. Serum inhibin B and anti-Müllerian hormone are not superior to follicle-stimulating hormone as predictors of the presence of sperm in testicular fine-needle aspiration in men with azoospermia. *Fertil Steril*. 2009;91(4):1279–84.
24. Pezzella A, Barbonetti A, D'Andrea S, Necozone S, Micillo A, Di Gregorio A, et al. Ultrasonographic caput epididymis diameter is reduced in

non-obstructive azoospermia compared with normozoospermia but is not predictive for successful sperm retrieval after TESE. *Hum Reprod.* 2014;29(7):1368–74.

25. Lotti F, Frizza F, Balercia G, Barbonetti A, Behre HM, Calogero AE, et al. The European Academy of Andrology (EAA) ultrasound study on healthy, fertile men: Scrotal ultrasound reference ranges and associations with clinical, seminal, and biochemical characteristics. *Andrology.* 2021;9(2):559–76.
26. Marconi M, Keudel A, Diemer T, Bergmann M, Steger K, Schuppe HC, et al. Combined trifocal and microsurgical testicular sperm extraction is the best technique for testicular sperm retrieval in low-chance nonobstructive azoospermia. *Eur Urol.* 2012;62(4):713–9.
27. Bouker A, Halouani L, Kharouf M, Latrous H, Makni M, Marrakchi O, et al. Step-by-step loupes-mTESE in non-obstructive azoospermic men, a retrospective study. *Basic Clin Androl.* 2019;29:11.

### **Publisher's note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.