Peer Review File

Article information: https://dx.doi.org/10.21037/atm-21-5100

Reviewer A

In this manuscript, the authors developed seminal plasma miRNAs as diagnosis markers and built a regression model to predict the m-TEST success rate.

Male infertility is a pressing problem for the human species. It is important to predict the m-TEST success for NOA patients. It is novel to use miRNAs in the seminal plasma as markers. However, the study does not begin with enough human subjects and the methods also lack sufficient scientific rigor.

Comment 1: The authors need to introduce the sources of seminal plasma and better explain their rational for using seminal plasma miRNAs as markers.

Reply 1: We appreciate for your valuable comment. Seminal plasma is the liquid component of semen which arises from secretions from the seminal vesicles (~65%), prostate (~25%), testes and epididymides (~10%) and periurethral glands (~1%) (Nat Rev Urol., 2014). Each tissue contributes specific molecular content, including proteins, circulating nucleic acids and lipids, which might indicate the status of the tissue. Secretions from testes and epididymides provide molecules related to spermatogenesis and maturation of spermatozoa (Hum Reprod., 2018). Clinical diagnosis of nonobstructive azoospermia, as well as the outcome prediction of micro-TESE, relies on the histological analysis of biopsied testicular tissue which might not accurately reflect the histology of testes owing to the focal spermatogenesis. Thus, seminal plasma, as the proximal fluid of testes might provide more accurate biomarkers. Several studies have reported proteins (ECM1 and TEX101), cell-free DNAs and RNAs as potential biomarkers, however, each biomarker has certain limitations (Sci Transl Med., 2013). For the protein markers, the diagnostic method relies on the method of ELISA which is less sensitive than qPCR-based method used in RNA biomarker test. As for cell-free DNA and mRNAs, they exist in seminal plasma in highly fragmented forms which make them lacking specificity. Compared with DNA and mRNA fragments, miRNAs have the characteristics of integral and stable, which makes them more suitable as biomarkers.

Changes in the text: we have modified our text as advised (see page 1-2).

Comment 2: Given the heterogeneity of NOA etiology, 18 sample sizes are insufficient.

Reply 2: Thank you for your rigorous consideration. This was also our very first concern in the study design step. It is true that NOA has different histopathological subtypes including Sertoli-cell-only Syndrome, maturation arrest and hypospermatogenesis based on histological analysis of

the biopsied testicular tissue. However, the results of diagnostic testicular biopsy might not accurately reflect the distribution of focal spermatogenesis in testis. Testicular sperm could also be found from testes of Sertoli-cell-only syndrome patients. It is our aim to establish a method to predict the existence of spermatogenesis foci and the possibility of sperm retrieval in micro-TESE instead of dividing histological subtypes. Thus, we divided samples based on the clinical micro-TESE results (the distribution of spermatogenesis foci under the surgical microscope) instead of the histological results. In this way, all enrolled samples in each group have homogeneous phenotype and six samples in each group for miRNA-seq is acceptable. For this study, the most important part is the establishment of the panel with candidate miRNAs, which means the sample size of the training set is pivotal. More samples will provide more accurate model. Thus, according to your suggestion, we added 20 more samples in training set for model improvement. We re-built the model with 76 samples by ordered logistic regression and test the accuracy of the model. The parameters of the model changed and we re-test the model with validation data. We found that the output of the model did not change comparing with the model built with 56 samples, and the predictive performance of the model is acceptable (AUC=0.927). The reason why the re-built model still gave the same output might be because the phenotypes of the input samples varied greatly, and the model is sensitive enough for distinguishing these samples.

Change in the text: We re-built the model with 76 samples. (See page 8, line 19-22)

Comment 3: How do the authors normalize the miRNAs? UMI only corrects the PCR error and does not give absolute quantification. The samples have different amounts of total RNAs. Based on the current method description, the authors will not be able to correct the total RNA difference among NOA human subjects.

Reply 3: This is a very critical point, and we also considered in the study design step. We spiked in exactly same amount (25 fmol) of synthesized **cel-miR-39** as reference in seminal plasma samples before RNA isolation step for normalization.

Change in text: See page 4, line 20-22.

Minor points:

1. Page 2 Line 1 change "transcription" to "mRNA stability".

Reply 1: we have made the change (see page 2 line 6)

2. Page 2 Line8, do the authors suggest miRNAs exist outside of exosomes in the seminal plasma?

Reply 2: This is a very good point of view. MiRNAs do not exist independently in body fluid. Circulating miRNAs either exist within the exosomes or they are bound by Argonaute protein family, other RNA binding proteins and, sometimes, other RNA species to form a complex in body

fluid. There are studies about exosome miRNAs and circulating miRNAs, but there are still no studies about miRNAs outside exosomes. It might be difficult to purify this part of miRNAs due to the technique issue. In this study, we purified circulating miRNAs including both miRNAs inside and outside exosomes.

3. Figure 2 label is too small to be legible.

Reply 3: We feel sorry for the inconvenience. We have made the change as you suggested.

Change in the text: we separated the figure 2 into 2 figures (figure 2 and figure 3). We removed the label from figure 2 and uploaded the labeled microRNAs in supplementary tables.

4. The authors need to discuss the functions and potential sources of miRNAs they identified.

Reply 4: Thank you for the suggestion. Two previous works from our team found that miR-34c cluster is essential for normal spermatogenesis (J. Bao et al. J Biol Chem.,2012; J. Wu et al. PNAS, 2014.). In this study we also found miR-34c-3p showed significant expression differences in NOA seminal plasma comparing with normal group. Thus, we have started to make a further study of the functions of miRNAs we identified in this study, which have not been studied in male reproduction, including miR-3065, miR-4446 and miR-518e.

Reviewer B

The study reports the establishment of a model for predicting the success rate of microdissection testicular sperm extraction in non-obstructive azoospermia patients. The study has included the screening phase to identify miRNA as biomarkers, the training phase to establish the model using machine learning and finally, the validation phase to test the model. The use of miRNA biomarkers in seminal plasma to reflect the spermatogenesis status is not a new idea. Some of the findings in the literature appear to be contradictory to the author's claim but the discussion on these findings was lacking. Besides, the classification of NOA samples did not include the subpopulation that affect the success rate of TESE.

Comment 1: NOA include patients with Sertoli-cell only syndrome and those with spermatogenic defects at different stages of spermatogenesis. The chance of identifying a sperm in patients after micro-TESE procedures varied with the severity of the spermatogenic defects but generally failed in patients without any germ cells. The study has group the patients into good, fair and poor in terms of the TESE results. What would be the status of spermatogenesis in these patients?

Reply 1: Thank you for your rigorous consideration. This was also our very first concern in the study design step. It is true that NOA has different histopathological subtypes including Sertoli-cell-only Syndrome, maturation arrest and hypospermatogenesis based on histological analysis of the biopsied testicular tissue. However, the results of diagnostic testicular biopsy might not accurately reflect the distribution of spermatogenesis foci in testis. Testicular sperm could also be found from some patients with Sertoli-cell-only Syndrome (from our clinical records). According to your suggestion, we summarized the histopathological results of the patients with different micro-TESE outcome in the table below (added as supplementary data). We also added a figure of histology and micro-TESE outcome in supplementary data.

Table S1.

	Histopathological Diagnostic Result			
Micro-TESE Outcome (case number)	Sertoli-cell-only Syndrome	Maturation Arrest	Hypospermatogenesis	Normal
Good (18)	0	0	6	12
Fair (47)	3	38	6	0
Poor (57)	40	17	0	0

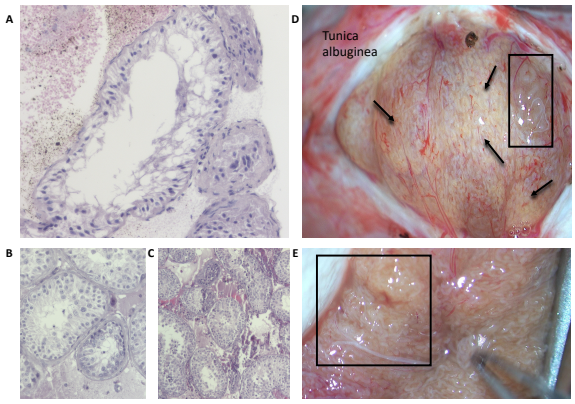


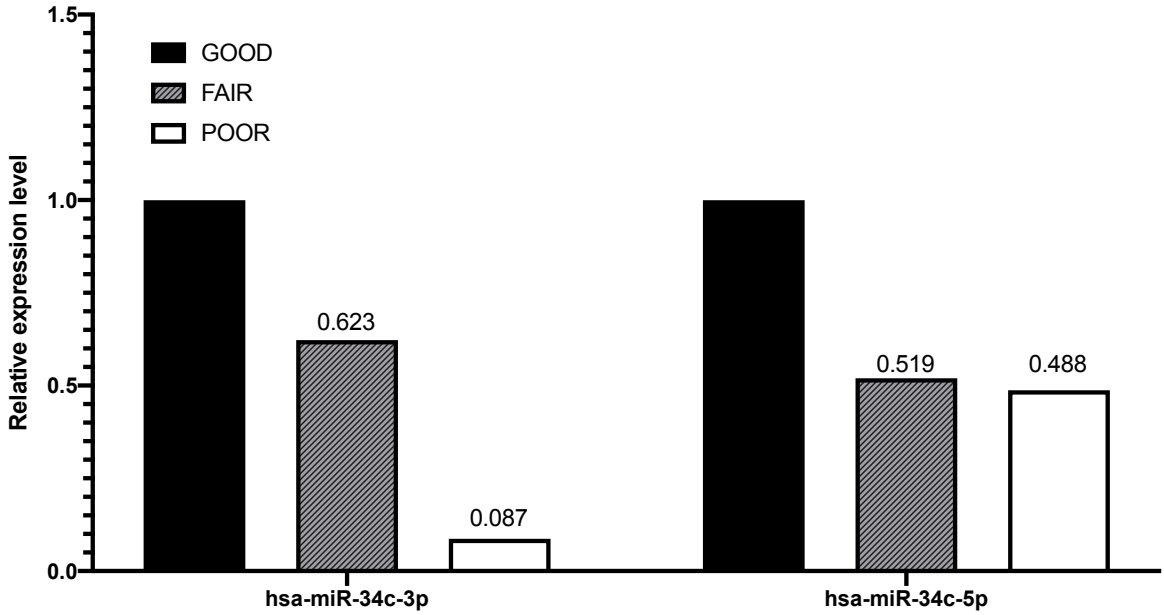
Fig.S4. **Histological diagnosis and micro-TESE outcome of enrolled patients**. (A) HE staining of Sertoli-cell-only Syndrome. (B) HE staining of Maturation Arrest. (C) HE staining of Hypospermatogenesis. (D) An isolated region of micro-TESE from the 'Fair' group, and the histopathological diagnosis was 'Sertoli-cell-only' Syndrome. The region in the black box shows tubules with normal spermatogenesis. The adjacent testicular tissue contains tubules with no spermatogenesis pointed by black arrows. (E) An isolated region of micro-TESE from the 'Good' group, and the histopathological diagnosis was 'Maturation Arrest'. The region in the black box shows tubules with normal spermatogenesis.

Change in the text: 1. We added table S1 in supplementary data; 2. We added figure S4 in supplementary data.

Comment 2: Following point 1, Zhang et al. have shown that the expression of miR-34c-5p was diminished in Sertoli cell-only syndrome patients and spermatogenic arrest groups (Fert Sertil 2021). In this study, miR-34c-3p was identified to be down-regulated in the "fair" group (fold change 0.62) and further decreased in the "poor" group (fold change 0.09). Since these two miRNAs are expressed from the same precursor, is there any difference observed in miR-34c-5p? The expression profiles appear to suggest the "fair" group to be the group with spermatogenic defects while the "poor" group correspond to the Sertoli-cell only syndrome. Echoing point 1, it would be necessary to validate and classify the patients with histology data.

Reply 2: Thanks for the valuable suggestion. Two previous works from our team found that miR-34c cluster is essential for normal spermatogenesis (J. Bao et al. J Biol Chem.,2012; J. Wu et al. PNAS, 2014.). Therefore, we particularly pulled out miR-34c from the sequencing data to check the expression patterns in each group. Both miR-34c-3p and miR-34c-5p showed significantly decreased expression pattern from the 'good' group to the 'poor' group (log2 FC= -6.43, *p* value= 2.15E-30). However, the fold change of miR-34c-5p was not significant from the 'fair' group to

the ‘poor’ group (log2 FC=0.85, *p* value=0.0021), which means miR-34c-5p might not be perfect candidate biomarker for distinguishing ‘fair’ samples from ‘poor’ samples based on our sequencing data. We added miR-34c-5p qPCR result as below. The reason for that might be that the histological diagnosis from testicular biopsy and the micro-TESE results are uncoupling in some cases, and this is also the reason why we screened seminal plasma biomarkers based on the micro-TESE outcome for improving the accuracy of preoperative examination.



Comment 3. Apart from miR-34c, the biomarkers identified in the literature appears to be different. How to explain the discrepancies? What would be the advantage of the model reported in this study as compared to those in the literature?

Reply 3: Some studies have reported similar method for biomarkers identification. The differences between our study and other studies are as below:

1. The aims of the studies are different, and the grouping methods are different. Some studies identified biomarkers from seminal plasma for histological subtypes identification, such as miR-201, miR-307 et al. Therefore, they collected and grouped samples based on histological diagnosis (Sertoli-cell-only syndrome, maturation arrest, hypospermiogenesis and obstructive azoospermia) and the sequencing results would be different from us.
2. The applicability of the diagnostic model/panel in clinic. Some studies identified biomarkers and established panel by logistic regression, and the output results for this panel would be binary (yes or no). The model we built was performed by ordered logistic

regression which means we made a rating evaluation for predict the possibility of the micro-TESE output. The output from our model gave three potential outcomes (good, fair and poor), which means the binary outputs could be further subdivided into three and become more robust.

Comment 4: Any information on the outcome of assisted reproduction after micro-TESE?

Reply 4: If testicular spermatozoa could be retrieved by micro-TESE, NOA patient could achieve fatherhood by performing intracytoplasmic sperm injection (ICSI). The time of sperm retrieval and the time of egg retrieval from the partner might not the same. There are three ways to choose for ICSI: egg and sperm retrieval in the same cycle, frozen egg recovery combined with sperm retrieval, and frozen sperm recovery combined with egg retrieval. Each way has certain limitations, such as the recovery rate of frozen eggs and sperms. In this study, according to the clinical records, among the 53 NOA patients who got successful sperm retrieval, 24 were received ICSI and 17 got successful pregnancy. The other 29 patients have frozen down the sperms.

Change in the text: we added description in 'Discussion' (see page 10-11).

Comment 5. The size of the labels in Fig 2 and 4 has made the figure unreadable.

Reply 5: We feel sorry for the inconvenience brought to the reviewer. We have made changes of the labels.