Peer Review File

Article Information: https://dx.doi.org/10.21037/tau-21-320

Response to Reviewer A

Comment 1： Line 69 may need to be reworded that it occurs in 10% of men presenting for fertility evaluations as clearly all men with NOA are infertile.

Reply 1: Thanks very much for your meritorious suggestion.

We are sorry for our misrepresentation.

According to your suggestion, we **reword this sentence more clearly**.

Changes in the text: line 70. For your convenience, the revised title is shown below:

Human infertility is a healthcare problem that has a worldwide impact (1). Currently, male factor infertility is increasing and accounts for about half of all infertility cases (2, 3). Non-obstructive azoospermia (NOA) is featured by entirely spermatozoa deficiency in semen (4, 5). Approximately 1% of all men and 10% of infertile male population are diagnosed as NOA (6). NOA is a multifactorial disorder and presents with diverse phenotypes (6-8). However, up to now, its molecular pathogenesis remains undefined.

Comment 2： I would recommend revision by a primary English writer for ease of reading and grammatical improvements.

Reply 2: Thanks very much for your valuable suggestion.

According to your professional comment, we invite Yixin Lee, **a native speaker, to polish our paper.** The revised text is highlighted in green color in the update manuscript. We also acknowledge his assistance in the "Acknowledgement" section.

Changes in the text: None.

Comment 3： The methodology is confusing. If this is a retrospective study why were consents signed to be in the study by patients? How as it double blinded if it was retrospective?

Reply 3: Thanks very much for your valuable advice.

We are sorry for our misrepresentation.

Actually, the written informed consent forms are not used to refer in particular to our study. All the patients attending the infertility clinic in reproductive center of Northwest Women's and Children's Hospital are informed that their biological materials may be used in bio-experiments, including our study and some other studies. Written informed consent forms were signed by every patient.

We marvel at your professionalism. In our study, the double blinded means that the researchers and the data analysts are in blind. This is not the typical double blinded experiment which researchers and subjects are in blind.

We are sorry again for our misrepresentation, and according to your suggestion, we **delete the "retrospective study"** and **explain the "double-blind"** in the manuscript.

Changes in the text: line 98. For your convenience, the revised title is shown below:

Subjects

In the present research, both researchers and the data analysts are in blind. A consecutive series of NOA subjects were selected from patients attending the infertility clinic in reproductive center of Northwest Women's and Children's Hospital who had a history of infertility of ≥12 months. Written informed consent forms were signed by every subject. Three times semen analyses were conducted after 3-7 days abstinence. Patients with history of chronic diseases, karyotype abnormalities, microdeletions of AZF region on Y chromosome and pelvic/spinal injuries were excluded. Testicular tissues were obtained from patients referred for micro-TESE.

Comment 4： The classification of MA needs to be subcategorized as early vs late MA depending on being prior to after meiotic division when the arrest occurs. These are very differing severities of histopathological architecture and should be looked at separately.

Reply 4: Thanks very much for your professional suggestion.

Early and late MA are very differing severities of histopathological architecture and should be looked at separately. We are so sorry that **only 7 MA patients are recruited** in our study, and it's difficult to separate them into early or late severities.

We have added this shortcoming into the limitation in the discussion part.

Changes in the text: line285-288. For your convenience, the revised title is shown below:

Other than the retrospective nature of this study, some nonnegligible limitations exist. Firstly, early vs late MA depends on being prior to after meiotic division when the

arrest occurs. These are very differing severities of histopathological architecture and should be looked at separately. The present study is limited by the small sample size (7 MA patients). Further study with more MA patients is in needed. Secondly, the subjects of our study are limited in our hospital. Further research with larger samples from other hospitals or centers are expected to validate our model. Thirdly, the representation of the population is not perfect, which is a common limitation in develop prediction models.

Comment 5： I am not sure the statement in the discussion that Beclin 1 expession is has obvious correlations with gonadotropins as these differences may be due to the severity of histopathology rather than Beclin 1 expression, this is a bit of extrapolation. The statement should be limited to Beclin 1 having higher expression in men with more severe histopathology which correlates to the clinical features described.

Reply 5: Thanks so much for your professional suggestion and we admire you for your rigorous thought.

According to your suggestion, we **revise the discussion** of the association between Beclin-1 and these clinical features **in a more rigorous description.**

Changes in the text: line 249-250 & line 256-265. For your convenience, the revised title is shown below:

The present study found that Beclin-1 were aberrantly expressed in three different pathological types of NOA, it was significantly up-regulated in SCOS when compared with HS or MA, meaning that Beclin-1 has a higher expression in men with more severe histopathology. Beclin-1 is a vital component of PtdIns3 kinase which stirs up autophagy (28). Beclin-1 complex (includes Beclin-1, VPS34, VPS15 and ATG14) also

plays the essential roles in autophagy initiation and regulation (14). These interesting results drove us to focus on Beclin-1 related autophagy is valuable in spermatogenesis. This view was consistent with other authors (11, 12). Thus, we have reasons to believe that autophagy-related gene Beclin-1 might be very helpful to guide personalized diagnosis and treatment of NOA. Moreover, Beclin-1 expression was also obviously positive correlation with serum LH, meanwhile significantly negative correlation with testicular volume, serum T, Johnsen score and pathologic type. Remarkably, more severe histopathology was reported to associated with higher level of serum LH (29), smaller testicular volume (30), lower level of serum T (31), and lower Johnsen score (32). Higher expression of Beclin-1 was also correlated with more severe histopathology. Hence, we are not sure that these associations between Beclin-1 and these clinical features may be due to the severity of histopathology rather than Beclin-1 expression. Further studies are expected to explore the role of Beclin-1 in regulating clinical features.

Comment 6： Multiple studies have established that SCO and early MA have poorer sperm retrieval rates than HS and late MA. That being said, the statement that Beclin 1 expression predicts SR rates cannot be made as SR rates are primarily associated with pathologic subtypes which as been well established. The statement should be limited to Beclin 1 expression being associated with the pathologic subtype rather than an independent predictor for SR.

Reply: Thanks so much for your precious suggestion.

Similar to the comment 5, our dicussion is a bit of extrapolation. According to your suggestion, we **revise the discussion** of the predictive value in sperm retrieval **in a more rigorous description.**

Changes in the text: line 276-283. For your convenience, the revised title is shown below:

Moreover, further multivariate analysis additionally disclosures that only Beclin-1 and Johnsen score have the most significantly predictive value of successful sperm retrieval. Beclin-1 was inversely related with successful sperm retrieval, however, Johnsen Score was directly related with successful sperm retrieval. As is known that sperm retrieval rates are primarily associated with pathologic types. Considering that pathologic types are ranked data, we did not add it into our multivariate analysis model. Our result indicated that Beclin-1 and Johnsen score, which had been reported to be related with pathologic types, showed significant predictive value in sperm retrieval. These findings further indicated that Beclin-1 might be vital biomarker for predicting successful sperm retrieval.

Response to Reviewer B

Comment 1： The authors have not considered the possibility that the relationship between hormonal parameters, Johnsen scores or Beclin expression and sperm retrieval are non-linear. Some authors have shown a U-shaped relationship in some of these results and sperm retrieval.

Reply 1: Thanks so much for your professional suggestion and we admire you for your rigorous thought.

Following your instruction, we screen the Pubmed and find some articles which show U-shaped relationships among hormonal parameters, Johnsen scores, and sperm retrieval. These research show that **U-shaped relationship is analyzed by Mann–Whitney U test in non-normal data**.

In our study, **our data were confirmed to be normal distributed** and the non-normally distributed data were transformed into normally distributed. We are sorry that the processes of confirmation and transformation were not added in our manuscript.

According to your suggestion, we add the processes of confirmation and transformation in the Materials and methods part.

Changes in the text: line 158-159. For your convenience, the revised title is shown below:

Statistical analysis

The non-normally distributed data were transformed into normally distributed using the cubed root transformation (18). First, the relative expressions of Beclin-1 among three different pathological types of NOA were evaluated using one-way analysis of variance (ANOVA). Thereafter, the correlations between the relative expression of Beclin-1and the clinical parameters such as age, testicular volume, serum hormone

levels (including LH, FSH, E2, T and PRL), and Johnsen score were estimated using Spearman's correlation test. Additionly, univariate and multivariable logistic regression analysis were performed to detect influence factors associated with successful sperm retrieval. Considering that pathologic types are ranked data, we score HS as 3, MA as 2, and SCOS as 1 in the present study. Finally, the receiver operating characteristic (ROC) curve analysis was conducted to calculate the area under the curve (AUC) and the optimal cutoff value (Youden's index) was predicted for the likelihood of successful sperm retrieval.

Comment 2： The surgical technique is not effectively described. Both the results reported as well as the surgical approach suggests that a limited sampling of testicular tissue was done for these patients rather than a more extensive microdissection evaluation of testicular sperm. This typically results in a direct inverse relationship between FSH and sperm retrieval chances, as reported in this article.

Reply 2: Thanks very much for your valuable suggestion.

According to your suggestion, we **add the detailed description of surgical technique** in our manuscript.

Changes in the text: line 117-125. For your convenience, the revised title is shown below:

Standard micro-TESE surgery was operated by experienced surgeons. Volume of each testicular was measured in the standard surgery room. Then, a longitudinal incision on the tunica albuginea of the testis was cut to reveal the seminiferous tubules under the professional operating microscope. Tubules with full appearing and opaque, which have possible sperm production, were gently dissected and placed in a petri dish. An

experienced embryologist dissected the seminiferous tubules and assessed the presence of sperm with the help of expert microscope. The positive result was confirmed when at least one sperm was found. Thereafter, a large fragment of testicular tissue of about $8 \times 4 \times 3$ mm^3 was cut out, no matter whether the sperm was harvested successfully.

Comment 3: The authors should be more clear about the relationship between results and sperm retrieval chances (i.e., direct relationship or inverse relationship.)

Reply 3: Thanks very much for your meritorious suggestion.

According to your advice, we **revise the sentences which describe the relationship more clear**.

Changes in the text: line 232-233. For your convenience, the revised title is shown below:

The results firstly demonstrated that Beclin-1 are aberrantly expressed in three different pathological types of NOA, it was significantly up-regulated in SCOS. Moreover, its expression was obviously positive correlation with serum LH, meanwhile significantly negative correlation with testicular volume, serum T, Johnsen score and pathologic type. Furthermore, a multivariate analysis demonstrated that Beclin-1 showed inverse relationship with testicular sperm retrieval. The cutoff value was 0.428, indicating that males, whose relative expression of Beclin-1 in testis was no more than 0.428, will have positive result in sperm retrieval in standard micro-TESE surgery. Our work has laid the foundations to develop Beclin-1 as a future molecular biomarker for diagnosis, predict and treatment of NOA.

Comment 4： Substantial English rewriting is needed.

Reply 4: Thanks very much for your valuable suggestion.

According to your professional comment, we invite Yixin Lee, **a native speaker, to polish our paper**. The revised text is highlighted in green color in the update manuscript. We also acknowledge his assistance in the "Acknowledgement" section.

Changes in the text: None

Response to Reviewer C

Comment 1： abstract: the 1, 2, 3 path scoring system can be removed from the abstract and included in methods of the manuscript instead. not relevant to the abstract

consider a few words on what Beclin 1 is in the first line of abstract ("Beclin-1 is an autophagy gene and higher levels suggest mammalian testicular damage")

Reply 1: Thanks very much for your meritorious suggestion.

According to your suggestion, we **remove the 1, 2, 3 path scoring system from the abstract** and **included it in methods of the manuscript instead**. Meanwhile, we **add the brief introduction of Beclin-1** in the abstract.

Changes in the text: line 37 & line 53 & line 166-167. For your convenience, the revised title is shown below:

Background: Beclin-1 is an autophagy gene and higher levels suggest mammalian testicular damage**.** The current study is to clarify the possible role of Beclin-1 in non-obstructive azoospermia (NOA) patients and explore its predictive value for testicular sperm retrieval.

Results: Our results showed that Beclin-1 are aberrantly expressed in three different pathological types of NOA, it was significantly up-regulated in Sertoli cell-only syndrome (SCOS) when compared with hypospermatogenesis (HS)(*P*=0.002) or maturation arrest (MA)(*P*=0.049). Moreover, Beclin-1 expression was obviously positive related with serum LH (rho=0.269, *P*=0.036), meanwhile significantly negative correlation with testicular volume (rho=-0.370, *P*=0.003), serum T (rho=-0.326, *P*=0.010), Johnsen score (rho=-0.318, *P*=0.012), and pathologic type (rho=-0.452, *P*<0.001). Furthermore, a logistic regression model demonstrated that Beclin-1 is an

important predictor of failed sperm retrieval (OR = 0.001, *P* = 0.007), which exhibited a pretty AUC=78.6 (*P* = 0.001).

Statistical analysis

The non-normally distributed data were transformed into normally distributed using the cubed root transformation (18). First, the relative expressions of Beclin-1 among three different pathological types of NOA were evaluated using one-way analysis of variance (ANOVA). Thereafter, the correlations between the relative expression of Beclin-1and the clinical parameters such as age, testicular volume, serum hormone levels (including LH, FSH, E2, T and PRL), and Johnsen score were estimated using Spearman's correlation test. Additionly, univariate and multivariable logistic regression analysis were performed to detect influence factors associated with successful sperm retrieval. Considering that pathologic types are ranked data, we score HS as 3, MA as 2, and SCOS as 1 in the present study. Finally, the receiver operating characteristic (ROC) curve analysis was conducted to calculate the area under the curve (AUC) and the optimal cutoff value (Youden's index) was predicted for the likelihood of successful sperm retrieval.

Comment 2： INTRO: in your 1st line on Beclin paragraph, it may be helpful to frame the role of beclin quickly or risk losing your reader. see above suggestion

Reply 2: Thanks very much for your valuable suggestion.

According to your suggestion, we **frame the role of beclin-1 quickly** in the 1st line on Beclin-1 paragraph in Introduction part.

Changes in the text: line 86-88. For your convenience, the revised title is shown below:

Beclin-1, located at chromosome 17q21 in humans, is the mammalian homolog of the yeast Autophagy-related gene 6 (Atg6) (10). Autophagy has recently been reported to be involved in spermatogenesis and fertilization (11-13). Higher levels of Beclin-1, one of the important autophagy related genes (14), is reported to suggest mammalian testicular damage. For instance, Huang et al indicated Aflatoxin B1 induced testicular damage and promoted autophagy by increasing the expression of Beclin-1, LC3 and Atg5 in mice testis (15). The results of Wang et al showed that cadmium accumulates caused testicular injury by increasing autophagy-related protein Beclin1 and LC3 (16). However, so far, the associations between Beclin-1 and human male infertility remain unknown. Hence, this study was conducted to find some track.

Comment 3： METHODS: please clarify how the study is retrospective and "in a double blind".

Reply 3: Thanks very much for your valuable advice.

We are sorry for our misrepresentation.

Actually, the written informed consent forms are not used to refer in particular to our study. All the patients attending the infertility clinic in reproductive center of Northwest Women's and Children's Hospital are informed that their biological materials may be used in bio-experiments, including our study and some other studies. Written informed consent forms were signed by every patient.

We marvel at your professionalism. In our study, the double blinded means that the researchers and the data analysts are in blind. This is not the typical double blinded experiment which researchers and subjects are in blind.

We are sorry again for our misrepresentation, and according to your suggestion, we **delete the "retrospective study"** and **explain the "double-blind"** in the manuscript.

Changes in the text: line 98. For your convenience, the revised title is shown below:

Subjects

In the present research, both researchers and the data analysts are in blind. A consecutive series of NOA subjects were selected from patients attending the infertility clinic in reproductive center of Northwest Women's and Children's Hospital who had a history of infertility of ≥12 months. Written informed consent forms were signed by every subject. Three times semen analyses were conducted after 3-7 days abstinence. Patients with history of chronic diseases, karyotype abnormalities, microdeletions of AZF region on Y chromosome and pelvic/spinal injuries were excluded. Testicular tissues were obtained from patients referred for micro-TESE.

Comment 4： RESULTS: please comment further on the OR of 0.001 for beclin. this seems extremely low, though it is significant in the model, what does it mean in the clinical context?

I did not see the sperm retrieval rates by pathology anywhere. this should be included since you are postulating that beclin expression is related to sperm retrieval success.

any references to beclin 1 levels in fertile male controls? is this a future direction worth mentioning?

Reply 4: Thanks very much for your precious suggestion.

The original data of the expression of Beclin-1 was non-normally distributed and we transformed them into normally distributed using cubed root transformation. This may lead to the little value of OR=0.001.

According to the sperm retrieval rates by pathology. Considering that pathologic types are ranked data, we did not add it into our multivariate analysis model. In our analysis, we add Johnsen score multivariate analysis model. Johnsen score is a series of continuous variable, and similar to pathologic types, it also has ability to reflect the spermatogenesis. Hence, Johnsen score is one of the pathological factor to predict sperm retrieval.

As you mentioned, the beclin-1 levels in fertile male controls are necessary, and further study are in needed to explore the differential expression of beclin-1 between NOA patients and fertile males, which may lays a foundation for exploring the role and the molecular mechanisms of Beclin-1 in NOA. According to your suggestion, we add it in our limitation in the manuscript.

Changes in the text: line 292-295. For your convenience, the revised title is shown below:

Other than the retrospective nature of this study, some nonnegligible limitations exist. Firstly, early vs late MA depends on being prior to after meiotic division when the arrest occurs. These are very differing severities of histopathological architecture and should be looked at separately. The present study is limited by the small sample size (7 MA patients). Further study with more MA patients is in needed. Secondly, the subjects of our study are limited in our hospital. Further research with larger samples from other hospitals or centers are expected to validate our model. Thirdly, the representation of the population is not perfect, which is a common limitation in develop prediction models. Fourthly, the beclin-1 levels in fertile male controls are necessary, and further study are in needed to explore the differential expression of

beclin-1 between NOA patients and fertile males, which may lays a foundation for exploring the role and the molecular mechanisms of Beclin-1 in NOA.

Comment 5： limitations: this needs to be reviewed and revised. lack of smoking data is not a significant limitation despite the reference included. This should be removed. Authors mention that this was validated in an independent dataset in their hospital but this is nowhere in the methods. Perhaps they mean to suggest it needs to be validated, in which case I concur. what do the authors mean that the population is not perfect?

Reply 5: Thanks so much for your professional suggestion.

According to your advice, we **remove the limitation of the lack of smoking data.**

We are sorry for our **misrepresentation in the description of subject recruitment**. According to your advice, we revised this limitation.

Changes in the text: line 288-290. For your convenience, the revised title is shown below:

Other than the retrospective nature of this study, some nonnegligible limitations exist. Firstly, early vs late MA depends on being prior to after meiotic division when the arrest occurs. These are very differing severities of histopathological architecture and should be looked at separately. The present study is limited by the small sample size (7 MA patients). Further study with more MA patients is in needed. Secondly, the subjects of our study are limited in our hospital. Further research with larger samples from other hospitals or centers are expected to validate our model. Thirdly, the representation of the population is not perfect, which is a common limitation in

develop prediction models. Fourthly, the beclin-1 levels in fertile male controls are necessary, and further study are in needed to explore the differential expression of beclin-1 between NOA patients and fertile males, which may lays a foundation for exploring the role and the molecular mechanisms of Beclin-1 in NOA.