Peer Review File

Article Information: https://dx.doi.org/10.21037/tau-21-922

Responses to the comments and suggestions of reviewer A

Comment 1: "How was disease recovery defined? Defined by a negative test following the last positive test? How often were COVID tests done to confirm a negative result?**"**

Reply 1: Thank you for the comment. The disease recovery of COVID-19 refers to the condition with lessened symptoms and two continuous negative SARS-CoV-2 nucleic acid tests. Specifically, the time point of disease recovery was referred to the last positive test time prior to getting lessened symptoms and two continuous negative tests or more. And the time interval of two continuous negative tests was required for more than 24 hours. Therefore, two or more continuous negative test of pharyngeal swab with more than 24 hours was performed to confirm a negative result.

Changes in the text: We have added some descriptions about the detailed definition of disease recovery in the method section (see page 7-8 lines 154-157).

Comment 2: "Please clarify how the follow up interval semen specimens were determined. If the cycle of spermatogenesis is approximately 3 months, the breakdown of recovery times of 120-150 days, 150-180 days and 180-210 days seems somewhat arbitrary. This would have been strengthened if participants had provided samples at each of these time points rather than just providing one additional specimen."

Reply 2: As described in our previous research, COVID-19 patients presented

a lower sperm quality after a recovery time of more than 90 days compared with patients who recovered from less than 90 days. It seems likely that COVID-19 patients had impairment of sperm quality within a recovery time of 90 days and manifested the damage after recovering from more than 90 days as a result of the cycle of spermatogenesis. Therefore, the impairment of sperm quality would possibly occur in any day of 90-180 days when spermatogenesis was impaired within 90 days of recovery time. Considering the sample size and the follow-up time and subjects' compliance, we analyzed the semen quality at a follow-up interval of 30 days.

As for the comment “This would have been strengthened if participants had provided samples at each of these time points rather than just providing one additional specimen”, we totally approve of it. It was designed that patients could come to the hospital to collect semen at various time points. However, it was a pity that it could not be carried out as planned due to the changes of the epidemic situation at that time and patients' own reasons.

Changes in the text: we have clarified how the follow-up interval semen specimens were determined in the discussion part (see page 14-15 lines 308-312).

Comment 3: “Please further clarify in Table 2 the timing of SA collection in the “recovered” patients.”

Reply 3: Thanks for your valuable comment. The median (interquartile range) timing of SA collection in the “recovered” patients was 177.5 (150.8-187.0) days. And we have clarified it in Table 2.

Changes in the text: we have added a row of recovery time information in Table 2 (see page 24 line 479).

Comment 4: “Figure 1 shows less change in total sperm within 90 days versus significant declines later, which is contrary to existing literature. Please discuss further.”

Reply 4: Thanks for your valuable suggestion. We have discussed it further in the discussion part.

Changes in the text: we have discussed it further in the discussion part (see Page 16 lines 331-335).

Responses to the comments and suggestions of reviewer B

Comment 1: “For reference #10 you state certain receptors are present in male reproductive tissues. Please be more specific as to which tissues.”

Reply 1: Thank you for the suggestion. As for male reproductive tissues, ACE2 receptor is expressed in the testes and seminal vesicles, while TMPRSS2 receptor could be detected in the prostate gland, testes and epididymis.

Changes in the text: We listed the specific male reproductive tissues in the introduction part and cited a corresponding literature (see page 5 lines 98-100).

Comment 2: “It would be helpful to provide some context as to how hospitalization of really any cause, flu, CHF, bacterial pneumonia affects sperm production and duration of impairment as it is known that acute illness has temporary effects on fertility.**”**

Reply 2: We deeply appreciate the comment. Acute infections such as

influenza viruses and pneumonia can have systemic effects on the body and have been found to affect semen quality as well. Decreased sperm motility and sperm count and changes in sperm morphology have been reported from 4 to 11 weeks after fever. The possible mechanisms are speculated as follows: a) fever causes increased testicular temperature and damages germ cell lines; B) Inducing orchitis and impairing the exocrine and endocrine function of testis. There is also evidence that flu may damage the DNA integrity of sperm (Sergerie M, et al., Fertil Steril 2007; Evenson DP, et al., J Androl 2000; Macleod J, Fertil Steril 1951).

Changes in the text: We have provided some context as to how acute infections such as influenza viruses and pneumonia affects sperm production and duration of impairment in the introduction and discussion section (see page 6 lines 123-127; page 12 lines 247-250).

Comment 3: **"**For reference 15-17 – how much decline in sperm quality (provide numbers) and over what time period.**"**

Reply 3: For reference 15, Ma et al. reported that 4 of 12 COVID-19 recovered patients with 78.5 days of median time between semen collection and disease onset had low sperm motility and higher sperm DNA fraction percentages. One of them could display about 16% declines after COVID-19 infection in total mobile sperm count. In reference 16, COVID-19 patients with a mean time of 25.5 days between the end of symptoms and semen collection, who recovered from moderate infection, had about 95% declines of total sperm number compared with controls. As for reference 17, COVID-19 recovered patients with 80 days of the median time between last positive pharyngeal swab test and semen collection had about 24% declines of total sperm number compared with age-matched healthy controls.

Changes in the text: we have added some detailed descriptions about how much declines of total sperm number and over what time period in the introduction section (see Page 5 lines 107-110; page 6 lines 111-120).

Comment 4: **"**Line 170 – control patient's without COVID OR (not and) preexisting infertility**"**

Reply 4: We deeply appreciate the comment. In the text, we have displaced the "and" with "or" in the text.

Changes in the text: We displace the "and" with "or" in the result part (see Page 10 line 199).

Comment 5: "Need to describe either in methods or results section the criteria for mild-moderate and severe COVID"

Reply 5: Mild type of COVID-19 referred to mild clinical symptoms and no radiological evidence of pneumonia on admission. Fever, respiratory symptoms or other symptoms, and radiological manifestations of pneumonia were considered as moderate COVID. Patients who met one of the following criteria were listed as severe cases of COVID-19: respiratory rate ≥ 30 /min; finger oxygen saturation at resting state ≤ 93%; arterial partial pressure of oxygen (PaO2)/fractional concentration of inspired oxygen (FiO2) ≤ 300 mmHg.

Changes in the text: We have added the criteria for mild, moderate and severe COVID-19 in the methods section according to the New Coronavirus Pneumonia Prevention and Control Program (7th ed.) published by the National Health Commission of China (see page 8 lines 162-170).

Comment 6: "The language regarding the 120-150 time period should be more careful. Although you acknowledge the limitation at the very end there are only 5 patients that were evaluated in that time period. Given the relatively large standard deviation of sperm concentrations, one would suspect with a larger sample size that this time period would NOT be significant given that g=<90 days vs >90 days recovery showed no difference. I think the conclusion that the 120-150 time period is significant is unfounded given the context of the data despite being technically statistically significant.**"**

Reply 6: Thank you very much for your suggestion, and we approve of your comment. We acknowledge the limitation of sample size in 120-150 time period for making a conclusion in the text and we have paid attention to the sentences and words in the manuscript.

Changes in the text: We acknowledge the limitation of sample size in the 120-150 time period for making a conclusion in the discussion and conclusion part (see Page 15-16 lines 328-331; page 17 lines 357-358).

Comment 7: "Again I think context should be provided to how impairment and recovery of spermatogenesis compare between COVID and other viruses and acute illness more generally."

Reply 7: As suggested, we have added some information about how the virus and acute illness affect sperm production and duration of impairment in our manuscript.

Changes in the text: we have added some information about how the virus and acute illness affect sperm production and duration of impairment in the discussion section (see page 11 lines 241-242; page 12 lines 243-250).

Responses to the comments and suggestions of reviewer C

Comment 1: “The 120-150 day bin of patients included only 5 patients. This is a very small sample number and makes this finding very difficult to assess.”

Reply 1: We totally approve of your comment. It certainly seems difficult to assess and make a conclusion using only five samples. On the other hand, it is likely that the actual level of total sperm number in the 120-150 days maintain as the level of 5 samples from the perspective of a small standard deviation of 5 patients’ total sperm number. Therefore, we acknowledge the limitation of sample size in the 120-150 time period for making a conclusion in the text, and propose that it needs a larger sample size in various time period to prove.

Changes in the text: we acknowledge the limitation of sample size in the 120-150 time period for making a conclusion and propose that it needs a larger sample size in various time periods to prove in the discussion and conclusion part (see page 15-16 lines 328-331; page 17 lines 357-358).

Comment 2: “How do the authors explain why a drop in sperm count occurred at 120-150 days post recovery, when the spermatogenesis cycle is 90 days? How to they explain this ‘apparent’ delay?**”**

Reply 2: As COVID-19 recovered patients at one or two months post-discharge presented persistent physical discomforts such as fatigue, dyspnea, chest pain and etc (D'Cruz RF, et al., ERJ Open Research 2021) (Carfi A, et al., JAMA 2020), it seems possible that the spermatogenesis is still affected one or two months after discharge. As a result of the spermatogenesis cycle of 90 days and the accumulative effect of COVID disease, the sperm quality could present the delay effect in the time period of 120-150 days.

Changes in the text: We have explained the delay effect of total sperm count in the discussion (see Page 15 lines 320-325).

Comment 3: “After the patients were binned by days of recovery, the authors only report total sperm count, when they do have the did collect data on sperm motility. Why was this not reported?”

Reply 3: Our previous investigation showed that total sperm number in the group of recovery time of ≥90 days was significantly lower than <90 days group while sperm motility had no significance between the two groups, and it revealed that only total sperm number in semen parameters had a delay effect (Ruan Y, et al., Andrology 2021). To explore the delay effect, we only analyzed the total sperm number. In addition, as suggested, we have made an analysis for sperm motility in the text.

Changes in the text: We make a box-plot figure to analyze the alteration of total sperm motility with recovery time at the end of the manuscript and add some related information in the result and discussion part (see Page 11 lines 232-236; page 15 lines 312-313; Figure 2).

Comment 4: “Why is the total normal motile sperm count not reported? What about morphology? It seems that this information is known based on how the study was conducted.”

Reply 4: Sperm morphology from COVID-19 recovered patients with a median recovery time of 80 days in the previous research was almost normal. Hence we just focus on the total sperm number, volume, sperm concentration, progressive motility and total motility to analyze the change of sperm quality. Thank you for your valuable comment. We will completely analyze the sperm

parameters in the future relevant research.

Changes in the text: none.

Comment 5: “Did the authors detect COVID in any of the semen samples in the recovery period?**”**

Reply 5: In our previous investigation, we have detected SARS-CoV-2 in the semen samples of COVID-19 recovered patients with a median recovery time of 80 days. Viral nucleic acid was not detected in 70 semen samples (Ruan Y, et al., Andrology 2021).

Changes in the text: We have added the testing information of COVID in the discussion section (see page 12-13 lines 264-265).