Peer Review File

Article Information: https://dx.doi.org/10.21037/tau-21-987

Reviewer A

In the current study, the authors retrospectively study their population of men being treated with clomiphene citrate and the corresponding changes in hormones and semen parameters. The manuscript adds value to the scant existing literature on this topic, but several points require clarification before it is suitable for publication:

Major

Comment 1: The study refers to excluding men with only one semen analysis, but upon further review of the manuscript, this seems to suggest excluding men without a pre- and post-treatment analysis. The lack of two semen analyses before and after treatment is a significant limitation and should be addressed.

Reply 1: We appreciate this request for clarification. Our study only included men with at least two semen analyses total: one prior to treatment with clomiphene and one after treatment. This has been the historic standard in reporting clomiphene data regarding efficacy, e.g.:

https://pubmed.ncbi.nlm.nih.gov/3918178/
https://pubmed.ncbi.nlm.nih.gov/1516979/

We agree that the data are limited by the lack of multiple pre- and post-treatment semen analyses, especially due to the wide variance in semen analysis results. We have updated our limitations section accordingly.

Changes in Text 1:
"Additionally, we were limited by the inherent wide variance in semen analysis results and therefore we observed wide variations in outcomes." has been added to the limitations section.

Comment 2: Why was the mean pre-treatment sperm concentration significantly above the WHO cutoff for normal sperm concentration?

Reply 2:
Thank you for bringing up this valid point. Our goal was to study how baseline sperm concentrations and gonadotropins would affect clomiphene's efficacy on total testosterone and semen analysis. Therefore, we included men who were above the WHO cutoff for normal sperm concentration to examine this question.

Changes in Text 2:

No additions to the text.

Comment 3: The pregnancy data, while adding some degree of value, contains little granular data (e.g. whether ART was required and how this rate might compare to a control population without treatment). The use of this data as justification for the FSH/LH cutoffs seems inappropriate.

Reply 3:

Our intention with the pregnancy data was to contextualize the cutoffs as we did not have a validation cohort to do this. We understand the reviewer's comments and have updated the table accordingly.

Changes in Text 3:

We have eliminated the pregnancy data from the study.

Minor

Comment 4: The abstract contains several grammatical mistakes and incomplete sentences and should be rewritten

Reply 4:

We thank the reviewer for the thorough review and have updated the abstract extensively.

Changes in Text 4:

Abstract is nearly rewritten.

Comment 5: Line 38 – the efficacy "remains" unclear

Reply 5:

We have rewritten the abstract and this error is corrected.

Changes in Text 5:

Abstract is nearly rewritten and the error is corrected.

Comment 6: The odds ratio for FSH and TT should be scaled to as to not use such arbitrarily small numbers (e.g. 1e-9)

Reply 6:

We appreciate the reviewer for suggesting another representation for our data. Our statistician had reviewed the literature available and found that scaling odds ratios by an arbitrary value could skew readers on the conclusions of our data.

https://www.feinberg.northwestern.edu/sites/firstdailylife/docs/resources-docs/jama.2018.norton.guidetostatisticsandmedicine.odds-ratioscurrent-best-practice-and-use.pdf

Changes in Text 6:

None

Comment 7: Line 62 – the assertion that azoospermic men do not benefit is in direct contradiction to the discussion section and should be reworded

Reply 7:

We thank the reviewer for pointing out areas where more distinction needs to be made. We have significantly re-constructed our discussion section to reflect a more nuanced appraisal of the effect of clomiphene on azoospermic men. We feel the manuscript is strengthened by this reinterpretation. Our results section and discussion have been updated.

Changes in Text 7:

Lines 152-156 now reads: "We found 34% (46/137) improved sperm concentration categorizations, 13% (18/137) worsened, and 53% (73/137) did not change. 23% (3/13) of azoospermic patients recovered sperm in the ejaculate (two patients with few sperm seen on centrifuged sample, one patient with approximately 100,000 sperm/mL)."

Lines 194-217 now reads: "Most patients did not improve WHO sperm concentration categorizations after taking CC; patients with non-obstructive azoospermia were least likely to benefit, although we did note that three out of 13 azoospermic patients achieved recovery of sperm in the ejaculate. It was unclear from our analysis whether this response is predictable based on *a priori* factors, due to the small sample size of the azoospermic subset. Hussein et al. found that 64.3% of men with non-obstructive azoospermia eventually recovered sperm in the ejaculate with prolonged clomiphene monotherapy, and sperm was retrieved with testicular sperm extraction in the other 35.7% of men (21). Other investigators have shown that optimizing TT in non-obstructive azoospermia using hormone therapy can improve the likelihood of finding sperm during microdissection testicular sperm extraction (mTESE) (22). Our data provide further context that may be helpful in managing expectations in men with azoospermia who wish to explore hormone treatment such as CC. Of note, we found that the degree of non-azoospermia did not affect the magnitude of improvement in sperm parameters when classifying as oligozoospermic or normozoospermic. These data suggest that non-azoospermic patients can experience similar magnitude of improvement when taking CC. It has been previously shown that oligozoospermic men experience benefit in sperm concentration and TT when prescribed CC (8). However, studies have not evaluated CC efficacy when comparing different baseline sperm concentrations."

Comment 8: Why were FSH and LH initial categorical cutoffs of 7 chosen a priori? Is there literature supporting this choice?

Reply 8:

We appreciate the reviewer for pointing out the categorical cutoffs. These cutoffs were chosen a priori as there is limited literature for a suitable gonadotropin cutoff. One of the goals of this paper was to establish suitable cutoffs to guide physicians on appropriate use of Clomiphene. We ran the same analysis using 6.5 and 7.5 as categorical cutoffs and resulted in similar trends. We have added in our methods section that these cutoffs were chosen a priori.

Changes in Text 8:

"Gonadotropin cutoffs were chosen arbitrarily as there is limited literature on cutoffs for the usage of CC" was added to the methods section.

Comment 9: Can you provide a reference for "mild oligozoospermia"? Ref 11 does not contain this cutoff, and neither do WHO published guidelines to my knowledge

Reply 9:

Mild, moderate, and severe oligospermia has been used in the infertility literature to classify degree of oligospermia.

Shaw's Textbook of Gynecology
https://pubmed.ncbi.nlm.nih.gov/20831771/

Changes in Text 9

Shaw's Textbook of Gynecology has been added to the reference list.

Comment 10: Sensitivity and specificity using the determined gonadotropin cutoff chosen via iterative t-testing method should be provided

Reply 10: The iterative t-testing methodology was proposed by our statistician as one option for evaluating threshold values for gonadotropins. We acknowledge the considerable limitations of this technique in our Discussion section. As this is not a robust "area-under-the-curve" type of analysis, we are unable to provide sensitivity or specificity with respect to this methodology.

Changes in Text 10:

None

Comment 11: Were all men with prior hormone therapies excluded? Specifically, men on remote TT, anastrozole, prior clomid therapy, etc or any duration or interval?

Reply 11:

All men with prior hormone therapies were excluded including men on remote TT, anastrozole, and prior clomiphene therapy (if there was no prior data) for any duration or interval. Our goal was to isolate the effects of clomiphene monotherapy and therefore wanted to minimize the effects that other medical therapies may have on the reported outcomes. In actuality, the effect of this exclusion criteria had minimal effect on the total of eligible patients.

Changes to text 11:

None

Comment 12: If parameters were not normally distributed, they should be presented as median (IQR) rather than mean/SD

Reply 12: We appreciate the reviewer's suggestions about reporting of these data. The statistics that we conducted for the data analysis would be more accurately conveyed if we report mean/SD. Additionally, there is significant previous literature on clomiphene where means/SD are suitable to represent these data.

https/pubmed.ncbi.nlm.nih.gov/22458540/
https/pubmed.ncbi.nlm.nih.gov/29873446/

Changes to the text 12:

None

Comment 13: How were men starting as normozoospermic treated in the WHO category improvement statistics?
Reply 13: These men were classified as no change as they did not change WHO categories of sperm concentration.

Comment 14: There are significant limitations in treating a continuous variable as a categorical variable (e.g. FSH, LH, BMI). This should be discussed in the limitation section.
Reply 14: BMI was treated as categorical since pre-defined cutoffs have been historically used to categorize individuals as normal weight, overweight, or obese. For gonadotropin analyses, we treated LH and FSH as continuous variables for multivariate analyses, and we used the ">7 versus <7" categorical cutoffs for separate analyses. We have provided a sentence in the limitations section to clarify this.
Changes to text 14: Lines 252-254 were added to the limitations section "We performed multiple subset analyses that treated gonadotropins as either continuous variables (multivariate analyses) or categorical variables (threshold analysis around an arbitrary cutoff of >7 or <7 miU/mL)."

Comment 15: Please provide the range of time interval between repeat semen analysis and clomid initiation. Was there a minimum amount of time required to be included?
Replay 15: The time range that was permissible for this study for follow up was first follow up less than 6 months as we historically use 3 month increments for SA tracking. The ranges have been added.
Changes in Text 15: The results section now reads "Median follow-up to second semen analysis was 3.8 months (range 1.2-5.7 months)."

Comment 16: Line 146 – should say "showed no improvement"
Reply 16: We have updated the manuscript accordingly.
Changes in Text 16: See above

Comment 17: I would argue that improving 3/13 (23%) of azoospermic men to cryptospermic or even severely oligozoospermic does indeed represent a significant improvement, as this may save these men from surgical sperm retrieval
Reply 17: We have significantly re-constructed our discussion section to reflect a more nuanced appraisal of the effect of clomiphene on azoospermic men. Thank you for illuminating the obverse perspective on these data. We feel the manuscript is strengthened by this reinterpretation.
Changes to text 17:
Lines 152-156 now reads: "We found 34% (46/137) improved sperm concentration categorizations, 13% (18/137) worsened, and 53% (73/137) did not change. 23% (3/13) of azoospermic patients recovered sperm in the ejaculate (two patients with few sperm seen on centrifuged sample, one patient with approximately 100,000 sperm/mL)."

Lines 194-217 now reads: “Most patients did not improve WHO sperm concentration categorizations after taking CC; patients with non-obstructive azoospermia were least likely to benefit, although we did note that three out of 13 azoospermic patients achieved recovery of sperm in the ejaculate. It was unclear from our analysis whether this response is predictable based on *a priori* factors, due to the small sample size of the azoospermic subset. Hussein et al. found that 64.3% of men with non-obstructive azoospermia eventually recovered sperm in the ejaculate with prolonged clomiphene monotherapy, and sperm was retrieved with testicular sperm extraction in the other 35.7% of men (21). Other investigators have shown that optimizing TT in non-obstructive azoospermia using hormone therapy can improve the likelihood of finding sperm during microdissection testicular sperm extraction (mTESE) (22). Our data provide further context that may be helpful in managing expectations in men with azoospermia who wish to explore hormone treatment such as CC. Of note, we found that the degree of non-azoospermia did not affect the magnitude of improvement in sperm parameters when classifying as oligozoospermic or normozoospermic. These data suggest that non-azoospermic patients can experience similar magnitude of improvement when taking CC. It has been previously shown that oligozoospermic men experience benefit in sperm concentration and TT when prescribed CC (8). However, studies have not evaluated CC efficacy when comparing different baseline sperm concentrations.”

Comment 18: Line 152-153 – why are these two improvements in TSC and TMSC listed as both favoring the lower LH cohort, but the values are presented in opposite order?
Reply 18: We thank the reviewer for pointing out this error in the ordering of our data. We have made the changes to reflect that improvements in TSC and TMSC both favor the lower LH cohort.
Changes to Text 18: The updated sentence now reads “When stratifying by pre-treatment LH (LH ≥ 7 vs LH < 7) miU/mL, we found a significant improvement in total sperm count (10.8 ± 41.1 M/mL vs 45.1 ± 125.6, p = 0.02) and total motile sperm count (19.1 ± 64.5 M vs 4.7 ± 17.7 , p = 0.04) favoring the LH < 7 miU/mL cohort.”

Comment 19: Line 153 – “we did not observe significant improvement in changes in TT and in sperm concentration”. – what cohort does this refer to? The entire group? Seems unlikely that men treated with clomid did not have an improvement in their TT
Replay 19: We agree with the reviewer that this sentence may be confusing. These changes in TT and sperm concentration refer specifically to the LH categorical data. We have updated the results section to clarify what this sentence refers to.
Changes to text 19: The sentence in question now reads “We did not observe significant improvement in changes in TT and in sperm concentration when using LH as a categorical variable (i.e., LH ≥ 7 versus LH < 7).”

Comment 20: How many of the couples who achieved a pregnancy required ART?
Reply 20: As previously discussed above, we have now eliminated the pregnancy data from the study.

Comment 21: Are the authors implying that men with elevated FSH should not be prescribed clomiphene? The clinical implications of these findings should be more carefully explained
Reply 21: Our results do not imply that clomiphene should not be prescribed to patients with elevated FSH. Our data only examines the relative benefit of clomiphene when stratifying by baseline parameters. Our findings suggest that men with elevated FSH exhibit less benefit with the drug than those who are hypogonadotropic, but further prospective studies would be required to fully characterize any diminishing returns from treating men with elevated FSH.

Comment 22: Figure 1 axes should be labeled
Reply 22: The figure axes have been re-labeled

Reviewer B

Comment 23: The authors have recognized the limitations within the current literature from a clinician's perspective in terms of predictive values for using Clomid. The paper builds off some prior work with an appropriate discussion of the limitations. The limitations are acknowledged and this provides some useful reference values.

Reply 23: We thank the positive insight from Reviewer B. We are happy that the limitations are appropriately acknowledged in the discussion.

Reviewer C

I congratulate the authors on their work assessing baseline gonadotropins differentiating hormonal and semen parameter response to CC.
This is a very well written manuscript overall, but I have some recommendations.
Comment 24: 1. In the Introduction the acronym for SERM is defined as a selective estrogen reuptake modulator, it should be a selective estrogen receptor modulator.
Reply 24: We thank the reviewer for the spotting the error. We have updated the manuscript accordingly.
Changes to text 24: **"**Selective Estrogen Reuptake Modulator" has been changed to "Selective Estrogen Receptor Modulator."

Comment 25: Although it is mentioned in the manuscript that there is not much guidance by ASRM/AUA guidelines, it may be worth including the specifics of the most recent guidelines: Clinicians may use SERMs for infertile men with low serum testosterone and Clinicians should inform the man with idiopathic infertility that the use of SERMs has limited benefits relative to results of ART.
Reply 25: We thank the reviewer's recommendation of including the specific guidelines that are included by the AUA. We agree this would highlight the importance of this study and have updated the manuscript accordingly.
Changes to text 25: Despite the frequent use of these medications, there is limited guidance

on the use of these treatment modalities in the American Urologic Association (AUA), European Association of Urology (EAU) and American Society of Reproductive Medicine (ASRM) guidelines (5, 6) which only recommends physicians to prescribe SERMs in patients with low testosterone for male infertility.

Comment 26: The dosages of CC 25 mg QD vs 50 mg QOD were included. Can rationale for use of one regimen over the other or when each was used be expanded on?
Reply 26: We thank the reviewer for pointing out the different prescription patterns that may exist for CC. The inclusion of both 25 mg QD and 50 mg QOD reflects the variation between providers and the time range of the included patient population. There were no discrepant variables that prompted the prescription of one dosing over another. Data remain limited on differences between these prescribing patterns, but Mazzola et al. in a 2013 AUA abstract suggested 25 mg QD may increase risk of tachyphylaxis. This limitation has been updated accordingly.
Changes to text 26:
None

Comment 27: In the discussion please define mTESE as microdissection testicular sperm extraction when first using this acronym.
Reply 27: We thank the reviewer for pointing out this detail. We have now included the definition of mTESE prior to first usage.
Changes to text 27: We have added “microdissection testicular sperm extraction” to the discussion.