**Author’s response to reviews**

**Title:** Assisted reproductive techniques with congenital hypogonadotropic hypogonadism patients: A systematic review and meta-analysis

**Authors:**

Yinjie Gao (1097156268@qq.com)

Bingqing Yu (113ybq@sina.com)

Jiangfeng Mao (maojiangfeng88@sina.com)

Xi Wang (wangxi8489@hotmail.com)

Min Nie (nm_pumch@aliyun.com)

Xueyan Wu (wsheyan@vip.sina.com)

**Version:** 4  **Date:** 24 Sep 2018

**Author’s response to reviews:**

Dear Prof. Esther Fagelson and Prof. Andrew A Dwyer:

Thank you for your letter and for the reviewer’s comments concerning our manuscript entitled “Assisted reproductive techniques improved fertility of congenital hypogonadotropic hypogonadism patients: A systematic review and meta-analysis” (ID: BEND-D-18-00229R3). Those reminders are valuable and very helpful for revising and correcting our paper. We have studied carefully and have made corrections which we hope meet with approval. The main corrections in the paper and responds are as follows:

1. Response to editor comments:

   (1) “We would also like to ask for you to provide more justification for the contributions of XYW and XW as currently they do not automatically qualify for authorship.”

   Response: We have re-written this part according to the editor’s suggestion. Every author’s contribution was listed in this portion. All authors have read and approved the manuscript, and ensure that this is the case.
(2) “Please upload your manuscript as a single, final, clean version that does not contain any tracked changes, comments, highlights, strikethroughs or text in different colors. All relevant tables/figures/additional files should also be clean versions. Figures (and additional files) should remain uploaded as separate files.”

Response: We have revised the manuscript and deleted all tracked changes, comments, highlights, strikethroughs or text in different colors. The figures and charts were adjusted to clean versions.

2. Response to reviewer reports:

(1) “CHH is a rare condition, and as such, data are limited and often are reported in only retrospective single study reports. I do not feel that the manuscript appropriately emphasizes this crucial point.”

Response: The research types were not restricted in the process of literature search. Retrospective or prospective studies, multicenter or single center studies all made sense to the meta-analysis. Unfortunately, for rare disease just as CHH, we have found only 20 retrospective studies consistent with the aims of the study. But we believed that 20 studies and 709 CHH patients could illustrate the issue to a large extent.

(2) “The description of CHH and incidence lacks precision and is inaccurate in some aspects. For instance, "secondary sexual immaturity and infertility" seems vague, why not simply state lack of puberty and infertility? GnRH is released "from" the hypothalamus not "in" the hypothalamus.”

Response: We have rechecked this part according to the reviewer’s suggestion and the following adjustments have been made:

① Congenital hypogonadotropic hypogonadism (CHH) is a disorder characterized by lacking of puberty and infertility, with low levels of circulating gonadotropins and sex steroids.  
② Two pathogenesis mechanisms exist for CHH. One is the reduced secretion of gonadotropin releasing hormone (GnRH) from the hypothalamus, and the other is the GnRH receptor defect in the pituitary.

(3) “The sexual discordance is not 5:1 (See Dzemaili et al Endocr Connect 2017). Incidence should be based on population studies - see original studies by Fromantin (French conscripts) and the Finnish population study by Raivio and colleagues. I question the 2/3 with anosmia (stated without reference). This is classically been about 50% in the literature.”

Response: We have consulted the references according to the reviewer’s suggestion and the following adjustments have been made: The incidence of CHH is approximately 1/86000–1/10000 (Fraietta R et al Clinics 2013), and the ratio of male versus female is about 3.6-1 which varies from race to race (Dzemaili S et al Endocrine Connections 2017). About 50% cases who
show anosmia/hyponosmia simultaneously called Kallmann syndrome (Kim SH et al Endocrinology and metabolism 2017).

(4) “There is no discussion of the variety of treatments for fertility inducing regimens.”

Response: The discussion of the treatments for CHH patients was elaborated in paragraph 2 of background. “Therapy for CHH depends on the patients’ desire for fertility at the time of treatment. Androgen replacement or sequential therapy of estrogen and progesterone can be used for patients who do not wish to have children. The combination of human chorionic gonadotropin (hCG) and human menopausal gonadotropin (hMG) is used to induce fertility. Pulsatile GnRH is another option for CHH patients who desire a pregnancy.”

(5) “Moreover, it can take 12-18 months to reach maximal testicular development and sperm count in males with CHH - should this not be noted and taken into consideration?”

Response: In discussion section, we reviewed a study published in 1997 which first emphasized that initiation of ICSI treatment after testicular maturity induced by hormonal treatment contributed to the success of ART (Yong EL et al Human reproduction 1997). Some other studies also stated that it can take 12-18 months to reach maximal testicular development and sperm count in males with CHH, so waiting may be advisable as maximal sperm counts are not attained until 12-18 months of treatment, and even longer in cases of cryptorchidism.

(6) “Please clarify how you determined patients CHH and not another form of hypogonadism and infertility. Presumably, patients with combined pituitary hormone deficiency could be included, as the treatment is the same”

Response: We have re-written this part according to the reviewer’s suggestion: All the 20 studies we included have clear patient inclusion criteria. The hypophyseal axis was checked by measuring TSH, cortisol, and prolactin, which showed no other combined pituitary hormone deficiency.

(7) “Please note that all studies were retrospective single centers, this is important as it has implications for heterogeneity and variability when interpreting the findings.”

Response: All included studies were retrospective, the integrity and consistency of the materials influenced the analysis of the outcomes and caused the deficiency of representation. As a limitation of our research, we explained in the discussion section. Although there were implications for heterogeneity and variability, the expanded sample size was still representative.

(8) “The description of pregnancy outcomes lacks precision i.e. duration was around 2 weeks in females and 6-12 months in males". This does not provide a rigorous description or interpretation of the data.”
Response: In the studies we included, the duration of ovulation induction in women before ART treatment was about 2 weeks, and gonadotropin therapy for spermatogenesis was 6-12 months in men. Male and female patients were treated during the same time respectively, so we meta-analyzed their fertility outcomes after ART treatment.

(9) “Similarly the description of ART approaches lacks precision and clarity (lines 176-177).”

Response: In the study we included, different types of ART treatment were involved, and specific descriptions of the different approaches were given in paragraph 3 of discussion section.

(10) “I do not believe that "ES" was defined (line 180).”

Response: We are so sorry to miss this point and have added the full term “effectiveness” to this position.

(11) “The I-square is reported as 73.06% yet no interpretation is offered. Please explain the significance of this in terms of heterogeneity and variability of studies. Please report confidence intervals rather than simple percentages, this would add to transparency and rigor.”

Response: We have re-written this part according to the reviewer’s suggestion: Heterogeneity between studies was assessed using the I² statistic with a cut-off of ≥50%, and the X² test with a p <0.10 to define a significant degree of heterogeneity. Where the degree of statistical heterogeneity was greater than this between trial results, possible explanations were investigated using subgroup analysis according to the gender. These were exploratory analyses only and may explain some of the observed variability, but the results should be interpreted with caution. The 95% confidence intervals were used to generate Forest plots of pooled relative risks for primary and secondary outcomes. We also performed a meta-regression analysis by age to show its influence on the pregnancy outcome.

(12) “What about description of weighted differences and subgroup analyses? If a qualitative interpretation was done this should be explicitly stated for transparency and clarity.”

Response: We have made supplementary remarks in the methods part: The sample size and risk of bias of included trials is crucial for the weighted differences, which means the more representative study is, the greater the weight is. Where the degree of statistical heterogeneity was greater than this between trial results, possible explanations were investigated using subgroup analysis according to the gender. It can tell the possible reason of heterogeneity.

(13) “It seem that not all studies reported adverse events, please note that this introduces a reporting bias and should be noted, similarly the discussion and conclusions should be softened to this end. The observation that there was no statistical difference (lines 199-200) is hindered by the fact that analysis is limited to the few studies that reported on this. This should be noted as a limitation.”
Response: We have made supplementary remarks in the limitation part: Last, it seems that not all studies reported adverse events, and some like OHSS is not considered to be an adverse event, so more studies should be included to avoid the reporting bias.

(14) “It is never noted in the discussion that ART is the only option for those who fail to conceive following fertility inducing hormone regimens. This may seem obvious, but it warrants mention.”

Response: We have pointed out that ART is a preferable option for CHH patients with unsuccessful long-term HRT in paragraph 1 of discussion section. It is necessary to mention in this meta-analysis.

(15) “Please specify forms of infertility (lines 208-209) rather than stating broadly other etiological infertility. Please report confidence intervals (see above).”

Response: We have re-written this part according to the reviewer’s suggestion: For this specific population, the overall pregnancy rate per ET cycle was about 46% (95% confidence interval 0.39 to 0.53), which was comparable to the patients with other etiological infertility including tubal factor infertility (TI), male factor infertility (MI) and unexplained infertility (UI).

(16) “There are numerous statements that are not backed up by the data, this may be acceptable in certain cases when summarizing publications in a review article, but given that a meta-analysis was performed, it seems logical that these data should be the basis for supporting conclusions.”

Response: CHH is a rare disease and the number of cases receiving ART treatment after HRT without pregnancy is not high enough. However, meta-analysis expands the sample size on the basis of existing studies, so that the results are more representative than separate research.

(17) “The methods exclude case studies yet the discussion includes reference 6, why is this included?”

Response: Reference 6 was not included in the 20 studies of this meta-analysis. As the first report of ART treatment in patients with Kallmann’ syndrome, it is of great significance to our review.

(18) “Please be precise, i.e. pregnancy rates were 50-60% (line 228) please refer to your analysis not those from the literature.”

Response: In this part, we reviewed some of the initial studies, and pregnancy rates from the large sample size studies were 50–60%. It reflects the general situation of ART treatment in CHH patients. And the results of our meta-analysis were showed in paragraph 1 of discussion section.
(19) “I disagree with the conclusion in lines 236-237. Waiting may be advisable as maximal sperm counts are not attained until 12-18 months of treatment (and even longer in cases of cryptorchidism).”

Response: We have re-written this part according to the reviewer’s suggestion: Waiting may be advisable as maximal sperm counts are not attained until 12-18 months of treatment, and even longer in cases of cryptorchidism. However, due to the difficulty of childbirth caused by age, the waiting time could not be extended indefinitely.

(20) “Line 238 seems overly descriptive i.e. "higher doses of gonadotropins are needed to induce ovulation" please base statements on your analysis rather than the literature.”

Response: Some studies we included stated that the ovaries were dormant and needed to be stimulated with higher doses of gonadotropins, which might theoretically increase the risk of OHSS in the CHH group. Hence, we evaluated this adverse event and explained that hormone therapy over a long period time could not bring better reproductive outcomes than ART treatment.

(21) “I am not sure the authors can definitively declare that ectopic pregnancy and OHSS are "common adverse events" given that not all reported these events.”

Response: We have made supplementary remarks in the limitation part: Last, it seems that not all studies reported adverse events, and some like OHSS is not considered to be an adverse event, so more studies should be included to avoid the reporting bias.

(22) “Please clearly state (in the limitations) that there are no prospective studies thus additional sources of potential bias.”

Response: We have pointed out that: Second, all included studies were retrospective. Hence, the integrity and consistency of the materials influenced the analysis of the outcomes and caused the deficiency of representation.

2. Response to minor concerns:

(1) “The manuscript could benefit from copy editing form a native English speaker. This would help clarity and precision.”

Response: According to the request, we have consulted native English speakers to polish the manuscript in English.

(2) “Abstract - please note that hormone replacement therapy to induce fertility is specialized i.e. combined or sequential gonadotropin therapy or pulsatile GnRH - this is an important
point and should be made explicitly clear. Also the comparison is not between treatments but rather outcomes to treatment. This point is important for precision.”

Response: We have re-written this part according to the reviewer’s suggestion in background section of abstract: After hormonal replacement therapy (HRT) including Androgen replacement or sequential therapy of estrogen and progesterone, The combination of human chorionic gonadotropin (hCG) and human menopausal gonadotropin (hMG) and Pulsatile GnRH, about 30% of patients with congenital hypogonadotropic hypogonadism (CHH) cannot produce sufficient gametes. Assisted reproductive techniques (ART) can efficiently treat infertility due to different causes.

(3) “Abstract - "efficiently" or effectively? Please be clear in these terms.”

Response: We have re-written this sentence according to the reviewer’s suggestion in background section of abstract: Assisted reproductive techniques (ART) can effectively treat infertility due to different causes.

(4) “Tables and figures should be revised. Table 1 is hard to read and lacks precision. For instance how is hCG, HMG, recFSH different from "gonadotropins"? "Control option" could simply be "control group", please use n=89 instead of 89 cases for clarity. Many open cells make it difficult to read, perhaps use n/a. I suggest using "outcomes" rather than "pregnancy outcomes" as this include implantation rate - I suggest reordering with live birth as the final outcome.”

Response: We have directly applied the contents of each study we included to analyze and summarize. For instance, the treatment and duration in Table 1, we described the gonadotropins treatment in a simple way, some studies told the detail such as “hCG” or “hMG”, but others only used “gonadotropins”. We have reordered table 1 according to the reviewer’s suggestion.

(5) “Table 2 - I suggest adding the maximum score (e.g. 24) to aid the reader in interpreting the quality score.”

Response: We have added the maximum score in the legend of table 2 according to the review’s suggestion.

(6) “Figure 4 - what about the high variability? Please address this.”

Response: We have added the legend below figure 2: Influence of age (X-axis) on pregnancy rate (Y-axis). The size of the circles indicates sample dimension.

(7) “Figure 5 is not readable due to resolution. Figure 6 is not readable.”

Response: We have readjusted the graphs for identification.

Once again, thank you very much for your comments and suggestions.
Best wishes,

Yinjie, Gao