PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Intracytoplasmic sperm injection (ICSI) versus conventional in vitro fertilization (IVF) in couples with non-severe male infertility (NSMI-ICSI): protocol for a multi-center randomized controlled trial
AUTHORS	Zheng, Danni; Zeng, Lin; Yang, Rui; Lian, Ying; Zhu, Yimin; Liang, Xiaoyan; Tang, Li; Wang, Huichun; Cao, Yunxia; Hao, Guimin; Liu, Jianqiao; Zhao, Junli; Wang, Rui; Mol, Ben; Rong, LI; Huang, He-Feng; Qiao, Jie

VERSION 1 – REVIEW

REVIEWER	Judy E. Stern, PhD Dartmouth-Hitchcock USA
REVIEW RETURNED	04-Apr-2019

GENERAL COMMENTS	This paper describes a protocol for a prospective, randomized, clinical trial to determine the outcomes of ICSI as compared with IVF cycles in couples with mild male factor infertility. This is a question for which an answer has long been sought and I commend the authors for attempting to answer this question in a systematic study. This is a large multicenter trail that will recruit 2,346 couples randomized to half receiving ICSI and half to IVF. This number was determined by a power analysis to define a difference of 7% between the groups. Inclusion criteria are on the basis of initial sperm count (5-15 million/ml) and motility (progressive motility10-32%). The endpoints will be pregnancy and live birth rates. The secondary outcomes will include adverse pregnancy outcome and birth outcomes. The study is well designed and the methods are clearly described. Recruitment has been ongoing since April of 2018 and will end in 2020. In general this is a well-designed study on an important topic. Nevertheless, I have a few questions about the methodology: • Has a power analysis been done to evaluate numbers needed for determining a difference in the secondary endpoints of birthweight, gestational age, infant congenital abnormality and so forth? Since these factors are determined on the basis of live births rather than cycles, the number needed to see a difference is likely to be greater than for the primary outcomes. If so, recruitment of additional couples might be advisable. • With regard to exclusion criteria: o Are there age limitations for the women or the men? o Are there any exclusions based on sperm morphology? In other words, is there any situation in which the sperm morphology is so poor that the couple would be excluded? o Nothing is said about female factors in these couples who may have multiple reasons for treatment. For example, are

	patients with diminished ovarian reserve excluded? If all patients, regardless of female factor are included, then groups either should have equal numbers of women with each diagnosis or the authors must adjust for diagnosis in their statistical analysis. Since no mention of female diagnosis is made in discussion of randomization it seems that all female diagnoses are included. The Methods section did not provide sufficient information for me to determine whether this will be considered in the statistical analysis of the outcomes. The authors mention patients with freeze-all strategies in the section of Methods on Blinding. They should clarify how the freeze all strategy fits into the overall protocol. Is this done for OHSS patients only? If not, when are these cycles used and how does their inclusion influence results? Frozen embryo transfer is also mentioned in the section on Embryo Transfer and Luteal Support. They should clarify the role, if any, of frozen cycles in contributing to overall live birth rates. It is unfortunate that the number of embryos transferred is “mostly limited to two...” For a study of this nature, the best course would have been to limit to transfer of one embryo only, but if that is not possible then at least they should limit it to two. Transfer of more than one complicates not only the live-birth and multiple birth rates, but also the analysis of the birth outcomes and may result in a need for recruitment of more patients to see differences in these. With regard to the secondary outcome of fertilization rate, the authors determine this as number ICSI per mature oocyte for ICSI and number per oocyte inseminated for IVF. This biases the numbers in favor of the ICSI group. Number of mature oocytes in the inseminated group can be determined on the day after insemination at the time of oocyte denudation and this number should be used as the denominator. Admittedly, the number that are mature is not as accurate as for the ICSI’d oocytes where denudation takes place on day of ovum pick-up since some oocytes will mature late. Nevertheless, completely immature oocytes should be subtracted from the denominator. More information is needed about the identification of congenital anomalies. Determinations of these rates differ depending on how defects are measured and which defects are included. It is not sufficient to say that all defects will be included since what is defined as a defect differs depending on recording method, timeframe and severity of defect. Perinatal death will be included as an outcome. Will neonatal death also be included? Under Data management the authors state that data collected on follow-up will be from a mixture of medical record and telephone interview. Since medical record information is generally more reliable than patient recall, it would help to have more information on the extent to which recall is to be relied on. This is particularly important with regard to complex factors such as congenital abnormalities that may not be accurately understood by parents.

REVIEWER	Ahmed Gibreel Mansoura University Egypt
REVIEW RETURNED	09-May-2019

GENERAL COMMENTS	This is a very well-written protocol with rigorous methodology. May only concern is related to the exclusion criteria where only women with no oocytes were excluded. I am a bit concerned if women is paying for their IVF or ICSI trial. In other words, if couples are paying for their IVF cycle then poor responders or low responders (women who produce less than 5 oocytes) should be excluded. As far as I am concerned, fertilization failure or poor fertilization may be expected in some of these couples with male factor subfertility hence we are by some way compromising their pursuit for a baby especially in low income country like china. If my understanding is wrong, and the study fund will cover the more than 2000 IVF cycle then it is quite fair and more comprehensive approach to recruit those couples to maximize the generalization of results

REVIEWER	Wellington Martins SEMEAR fertilidade
REVIEW RETURNED	10-May-2019

GENERAL COMMENTS	Authors should also report the cumulative live birth/ongoing pregnancy rate per oocyte retrieval, as this is a very important outcome for this trial. Not only the results of the first embryo transfer are relevant, if one method results in fewer good embryos, this will be better noticed if authors report the cumulative live birth rate. In my opinion, the cumulative live birth rate is more important than the live birth rate following the first embryo transfer for this trial. Eligibility criteria: Why not including sperm concentration > 15 millions/mL? Why not including progressive motility > 32%? It would be easier to use a single criterion; for example, a total motile count > 5 millions. Or, if restricting by the upper limit, a total motile count of 5-20 millions. If authors are to include only couple with slightly abnormal semen analysis (i.e not including couples with normal semen), this should be better defined in the abstract. The allocation concealment is not clearly reported. Authors need to provide more details regarding the IVF group. Authors need to detail: Total volume; Number of oocytes per well; Final motile sperm concentration (motile sperm/mL) . Outcomes: fertilization rate should be assessed only by oocyte retrieved to permit fair comparison. Only in the ICSI group it will be possible to assess fertilization rate per oocyte injected.

VERSION 1 – AUTHOR RESPONSE

Reviewer #1:

1. The reviewer asked whether a power analysis has been done to evaluate numbers needed for determining a difference in the secondary endpoints of birthweight, gestational age, infant congenital abnormality and so forth? Since these factors are determined on the basis of live births rather than cycles, the number needed to see a difference is likely to be greater than for the primary outcomes. If so, recruitment of additional couples might be advisable.

Thank you for your advice. The primary endpoint of this trial is live birth lead by ongoing pregnancy, which is the most important outcome for infertile patients and considered to be the "gold standard" index in studies evaluating assisted reproductive technology. Therefore, we calculated the sample size according live birth rate.

Birthweight, gestational age and infant congenital abnormality are secondary endpoints of this trial. They are important but we do not think we should increase the sample size based on these secondary outcomes. We have referred to this in the revised manuscript and added the sentence (please see page 4, line 1-2).

Please keep in mind that if we want to distinguish the difference between the two study arms on these indexes, we might need to recruit more couples. Since we do not know the outcome of the study, recruitment of more couples might expose more couples to a lower live birth rate.

Finally, meta-analysis also reported that there was no risk difference between children's birth defect conceived by IVF and/or ICSI (Wen, J., et al., Birth defects in children conceived by in vitro fertilization and intracytoplasmic sperm injection: a meta-analysis. Fertility and Sterility, 2012. PMID: 22480819). Therefore, we stick on calculating the sample size based the live birth rate.

2. The reviewer asked are there age limitations for the women or the men?

Thank you for this question. In the introduction of our manuscript, we have argued that ICSI was introduced successfully to treat male infertility initially and now the first-line therapy in severe male infertility. In our clinical practice, the decision regarding performing ICSI or conventional IVF is always based on sperm parameters and previous IVF/ICSI history, instead of couple's age. A recent study from Canada reported on the pregnancy outcome of ICSI and IVF in women with advanced age (Tannus, S., et al., The role of intracytoplasmic sperm injection in non-male factor infertility in advanced maternal age. Hum Reprod, 2017. PMID: 27852688). Results found that the use of ICSI for the sole indication of advanced maternal age showed no difference with conventional IVF. Therefore, we did not set the age limitation for couples in our trial to increase the generalizability of the trial findings.

3. The reviewer asked are there any exclusions based on sperm morphology? In other words, is there any situation in which the sperm morphology is so poor that the couple would be excluded?

Thank you for your comments. As we know, there are no widely accepted criteria for the treatment technique about isolated teratozoospermia. As a consequence, decisions for couples with isolated teratozoospermia are empirical.

We have searched the literature about this theme. It has been shown in some studies that patients with teratozoospermia achieved good fertilization in conventional IVF as long as the sperm concentration and motility were within the normal range according to WHO standards. A study from China also evaluated the effect of isolated teratozoospermia in ICSI or IVF cycle and found no differences for fertilization and pregnancy outcomes (Fan, W., et al., Outcome of conventional IVF and ICSI on sibling oocytes in the case of isolated teratozoospermia. J Assist Reprod Genet, 2012. PMID: 22733316).

While discussing our trial design, a limitation of sperm morphology (in other word exclusion of patient with teratozoospermia) was not found to be needed. However, we will evaluate sperm morphology on

the day of oocyte retrieval, which is a marker for the effectiveness of ICSI. Please see the added description in page 12, line 15-16.

4. The reviewer states that nothing is said about female factors in these couples who may have multiple reasons for treatment. For example, are patients with diminished ovarian reserve excluded? If all patients, regardless of female factor are included, then groups either should have equal numbers of women with each diagnosis or the authors must adjust for diagnosis in their statistical analysis. Since no mention of female diagnosis is made in discussion of randomization it seems that all female diagnoses are included. The Methods section did not provide sufficient information for me to determine whether this will be considered in the statistical analysis of the outcomes.
We appreciate your comment. To ensure ethicality, feasibility and generality of this large-scale multicenter randomized controlled study, our group did a lot literature review during proposal of it. Due to limited words, we just said that “many studies have indicated the routine use of ICSI in non-male factors infertility was not recommended to improve the clinical outcomes” in background section of our manuscript. Moreover, except studies focused on “non-male factor infertility”, several studies focused on just one female factor infertility, such as poor ovarian response, endometriosis, PCOS, anatomical problem, tubal factor infertility, as well as unexplained infertility. When restricted to different indications of non-male factor, pregnancy outcome didn’t have significant difference between two groups. Based on these literature results, we decided to enroll all kinds of female factors into our study population. Thank you for reminding as these factors should be adjusted into our statistical model. We have now mentioned that couples with various female indications for IVF will be included. Please see supplemented contents in page 10, line 7.

5. The reviewer thought authors should clarify how the freeze all strategy fits into the overall protocol. Is this done for OHSS patients only? If not, when are these cycles used and how does their inclusion influence results? Frozen embryo transfer is also mentioned in the section on Embryo Transfer and Luteal Support. They should clarify the role, if any, of frozen cycles in contributing to overall live birth rates.
Thank you for this question. In all our study centers, transfer of fresh embryos is the usual practice when fresh embryos are available. In some cases, all embryos may be cryopreserved without a fresh-embryo transfer, most commonly to prevent the ovarian hyperstimulation syndrome. In addition, freeze all strategy will be used when the following situations occurred in current fresh IVF/ICSI cycle: hydrosalpinx, elevated progesterone in hCG day, endometrial factors (endometrial polyps, endometrial cavity fluid and thin endometrium), systematic diseases (stomachache, fever or cold), and sudden accident of patients. We have also assumed that patient receive fresh or frozen embryo transfer cycle would be randomized equally into ICSI/IVF group with our 1:1 randomization design. Please see revised description of embryo transfer in page 13, line 7-13.
It now reads “Fresh or frozen-thawed embryo transfer will be decided by physicians according to conditions of patients. Transfer of fresh embryos is the usual practice when fresh embryos are available in all our study centres. In some cases, all embryos may be cryopreserved without a fresh-embryo transfer, most commonly to prevent the ovarian hyperstimulation syndrome. In addition, a freeze-all strategy will be used in the following scenarios: hydrosalpinx, elevated progesterone in hGC day, endometrial factors (endometrial polyps, endometrial cavity fluid and thin endometrium), systematic diseases (stomach-ache, fever or cold), and sudden accident of patients.”

6. The reviewer said it is unfortunate that the number of embryos transferred is “mostly limited to two...” For a study of this nature, the best course would have been to limit to transfer of one embryo only, but if that is not possible then at least they should limit it to two. Transfer of more than one complicates not only the live-birth and multiple birth rates, but also the analysis of the birth outcomes and may result in a need for recruitment of more patients to see differences in these.
Thanks for your composite. We used to have inapposite description for the number of embryos transferred. We have required all our study centers to transfer up to two best-quality embryos instead

of “mostly limited to two embryos transferred”. One embryo would be transferred if there were uterine malformation, history of uterine surgery or cesarean section, in order to reduce risks of multiple birth rates and adverse pregnancy outcomes. We revised our description in the section of “Embryo transfer and luteal support”, please see page 13, line 18-20.

7. The reviewer said with regard to the secondary outcome of fertilization rate, the authors determine this as number ICSI per mature oocyte for ICSI and number per oocyte inseminated for IVF. This biases the numbers in favor of the ICSI group. Number of mature oocytes in the inseminated group can be determined on the day after insemination at the time of oocyte denudation and this number should be used as the denominator. Admittedly, the number that are mature is not as accurate as for the ICSI’d oocytes where denudation takes place on day of ovum pick-up since some oocytes will mature late. Nevertheless, completely immature oocytes should be subtracted from the denominator.
Thank you so much for the correction and interpretation. Reviewer 3 also mentioned this question that only in the ICSI group it will be possible to assess fertilization rate per oocyte injected. To permit fair comparison, we can use number of oocytes retrieved as our denominator in per protocol analysis, in addition to intention-to-treat analysis. This would address your concern. Please see revised description in page 14, line 10-11.

8. The reviewer suggested that more information is needed about the identification of congenital anomalies. Determinations of these rates differ depending on how defects are measured and which defects are included. It is not sufficient to say that all defects will be included since what is defined as a defect differs depending on recording method, timeframe and severity of defect.
Thank you for your comment. We have discussed it and decided to focus on trisomy 13,18 and 21, congenital heart disease, cleft lip, excessive numbers of fingers or toes, and neural tube defect as these diseases are common congenital anomalies in China (Report on the Prevention and Treatment of Birth Defects in China (2012), NHFPC, 2012). We will still include ‘other’ anomalies, as that information (though not predefined) should not be neglected.
For recoding information about congenital anomaly, we can obtain related information from their obstetrical medical records. Diagnosis of congenital anomaly is recorded in the discharge information coded as the International Classification of Diseases, revision 10 (ICD-10) coding system. Through this we can know exactly about the related congenital anomaly. Please see our supplemented content in page 15, line 19-24.

9. The reviewer thought perinatal death will be included as an outcome. Will neonatal death also be included?
We thank you your suggestion. We have added neonatal death as a secondary outcome, as the definition was obtained from Zegers-Hochschild, F., et al., The International Glossary on Infertility and Fertility Care, 2017. Fertility and Sterility, 2017. Please see page 15, line 28.

10. The reviewer suggested that Under Data management the authors state that data collected on follow-up will be from a mixture of medical record and telephone interview. Since medical record information is generally more reliable than patient recall, it would help to have more information on the extent to which recall is to be relied on. This is particularly important with regard to complex factors such as congenital abnormalities that may not be accurately understood by parents.
Apology for the confusion. All the final outcomes reported in the trial will be extracted from medical records. We have adjusted the texts in page 15, line 32. It now reads “The data collected for the trial will be a mixture of routinely clinical data and information from follow-up, which are verifiable from the medical record.”

Reviewer #2:

1. The reviewer had a bit concerned if women is paying for their IVF or ICSI trial. In other words, if couples are paying for their IVF cycle then poor responders or low responders (women who produce less than 5 oocytes) should be excluded. As far as I am concerned, fertilization failure or poor fertilization may be expected in some of these couples with male factor subfertility hence we are by some way compromising their pursuit for a baby especially in low income country like china. If my understanding is wrong, and the study fund will cover the more than 2000 IVF cycle then it is quite fair and more comprehensive approach to recruit those couples to maximize the generalization of results.

Thanks for your comment. During the preliminary literature review of this project, we confirmed the efficiency of IVF/ICSI among patients with poor ovarian respond: several studies compared ICSI with conventional IVF in patients less than 3 oocytes retrieval and found that fertilization and pregnancy outcomes were comparable.

In our clinical practice, the fertilization method is chosen by embryologists according to semen analysis instead of oocyte number on the day of oocyte retrieval among all patients. For patients with poor ovarian response, our physicians explain possible risks, outcomes and other concerns before oocyte retrieval, independent whether they participate in our trial or not.

It is expected that patients with poor ovarian response in our trial will be equally randomized to ICSI or IVF in a 1:1 randomization design. Therefore, risk of fertilization failure or poor fertilization tend to be similar after randomization to ICSI or IVF group. Also, randomization would not affect further fertilization or pregnancy outcome for patient with poor ovarian respond.

In this study, the cost of IVF/ICSI is paid by the participants. While during the informed consent process, we always emphasize this important issue on costs and make sure the participants are fully informed. Only patients who are willing to afford the cost of IVF or ICSI cycles will sign the informed consent, for included study participation, we will offer them clinical consultation or help as much as possible.

Reviewer #3:

1. The reviewer thought authors should also report the cumulative live birth/ongoing pregnancy rate per oocyte retrieval, as this is a very important outcome for this trial. Not only the results of the first embryo transfer are relevant, if one method results in fewer good embryos, this will be better noticed if authors report the cumulative live birth rate. In my opinion, the cumulative live birth rate is more important than the live birth rate following the first embryo transfer for this trial.

Thank you for this comment. While cumulative live birth per oocyte retrieval is important, it does not take time into consideration. For instance, a couple having a live birth after 10 embryo transfer attempts will be weighted the same as a couple with a live birth after the first embryo transfer. Therefore, live birth rate after the first embryo transfer serves better as the primary outcome, incorporating both effectiveness and time. In addition, the findings in this trial cannot justify whether there is a causal inference between good quality embryos and cumulative live birth rate. Such a research question should be answered by a different study design. Nevertheless, number of good quality embryos has been included as one of the secondary outcomes, which seems sufficient to address the reviewer's interests in good quality embryos.

2. The reviewer asked why not including sperm concentration > 15 millions/mL? Why not including progressive motility > 32%? It would be easier to use a single criterion; for example, a total motile count > 5 millions. Or, if restricting by the upper limit, a total motile count of 5-20 millions. If authors are to include only couple with slightly abnormal semen analysis (i.e not including couples with normal semen), this should be better defined in the abstract.

Thanks for this question. ICSI was originally introduced as add-on to IVF for couples with severe male infertility, but is in current clinical practice also used in couples with mild male or even unexplained infertility.

As it is generally accepted that severe male infertility can be treated successfully with ICSI, we decided to focus on patient with slight or moderate oligospermia/asthenozoospermia. In 2010, World Health Organization (WHO) defined a normal quality semen sample as having a sperm concentration of 15 million/mL or greater, progressive motility (PR) 32% or greater (fifth edition of WHO laboratory manual for human semen). According this criterion, we defined our patients with sperm concentrate between 10-15×106/ml and/or sperm with progressive motility (type a+b) between 10-32% as "non-severe male infertility". We have adjusted the abstract. Please see the added description in page 3, line 9-10. It now reads "We will study couples with non-severe male infertility (defined as a semen concentrate 5-15×106/ml or sperm with a progressive motility 10-32%) scheduled for their first or second ICSI or IVF cycle."

3. The reviewer suggested the allocation concealment should be more clear.
Thank you for this question. On the day of oocyte retrieval, administrative staffs in the IVF laboratory will log into the trial system to randomize and allocate participants to receive either ICSI or IVF. Only embryologists know the allocation immediately. They will do the related insemination, specific allocation will be blinded to patients and clinicians until day of embryo transfer, and participants will know about their allocation on the day of embryo transfer, as well as clinicians who perform the embryo transfer. We have adjusted the manuscript. Please see revised content in in page 11, line 3-6. It now reads "On the day of oocyte retrieval, administrative staff in the IVF laboratory will log into the trial system to randomize and allocate participants to receive either ICSI or IVF. Initially, only embryologists will know the allocation. Participants and clinicians will be informed about the randomized allocation on the day of embryo transfer for participants with fresh embryo transfers and the day of embryo freezing for couples with freeze-all strategies. Prior to these dates, participants and clinicians will still be unaware of randomization allocation."

4. The reviewer suggested authors need to provide more details regarding the IVF group.
Authors need to detail: Total volume; Number of oocytes per well; Final motile sperm concentration (motile sperm/mL).
Thank you for your suggestion. In this multi-center study, the specific procedure in insemination and embryo culture should be adhered to routine from different study center. Due to word limit, we described the main text instead of describe all details in all centers. With the randomization stratified within study centers, different IVF procedures would not influence pregnancy outcomes of participants. In addition, we added some content related to culture of COCs after oocyte retrieval, please see page 12, line 5-7.

5. The reviewer suggested that fertilization rate should be assessed only by oocyte retrieved to permit fair comparison. Only in the ICSI group it will be possible to assess fertilization rate per oocyte injected.
Thank you for this correction. Reviewer 1 also mentioned about this. We revised the related description according your suggestion. Please see page 14, line 10-11.

VERSION 2 – REVIEW

REVIEWER	Judy E. Stern, PhD Dartmouth-Hitchcock USA
REVIEW RETURNED	16-Jul-2019

GENERAL COMMENTS	The authors have adequately addressed reviewer concerns.